ILLUMINATING THE DARK UNIVERSE WITH GALAXY CLUSTERS

by

Sarah Louise Mulroy

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Astrophysics and Space Research Group
School of Physics and Astronomy
The University of Birmingham
April 2017

Abstract

Structure in the Universe formed from tiny density perturbations that grew into the complex cosmic web we see today. At the nodes of this web we find galaxy clusters – the largest gravitationally bound objects ever observed. Composed of dark matter, gas and galaxies, they provide an environment within which to investigate a multitude of physical processes. Their abundance and their properties provide an insight into the evolution of the Universe, so they are also important probes of cosmology. Knowledge of their mass is critical for cosmology, but as most of this mass is in the form of dark matter it is a complex measurement.

To determine the mass of the large samples of clusters required, we instead focus on measuring another cluster observable and its scaling relation with cluster mass. Ideally these observables should be easy to measure from upcoming surveys that will image large areas of the sky. The scaling relation with mass must be well constrained, and preferably with minimal intrinsic scatter.

In this thesis I use a sample of low redshift galaxy clusters with high quality multiwavelength observations to investigate these cluster observables and their scaling relations with mass.

I introduce the near-infrared luminosity of a galaxy cluster, an observable easily measured from shallow survey data. As a measure of the stellar mass of a cluster, we expect near-infrared luminosity to scale tightly with mass, and I show that this is indeed the case.

Expanding this work to optical luminosities by looking at bluer wavelengths, I find that the low intrinsic scatter is not limited to near-infrared wavelengths. This demonstrates that cluster luminosity in a range of wavelengths is a promising observable for cluster cosmology. With these measurements, I next investigate the trend of the total colour of galaxy clusters with indicators of the level of disturbance in the cluster, finding that the more disturbed clusters have a smaller range of total cluster colours than the less disturbed clusters. This suggests a smaller range in the star formation rate in disturbed clusters, perhaps caused by a standardising effect of major merger activity.

Finally, I investigate the scaling relations of cluster observables from across the electromagnetic spectrum, using a new maximum likelihood method. This method corrects for selection effects and provides tight constraints on both the scaling relation parameters and the covariances between observables at fixed mass. The results are consistent with self-similarity, in which clusters are formed from a single spherical collapse driven by gravity, and with a closed box picture, in which clusters maintain their baryon budget.

Acknowledgements

A huge thank you to my parents, who may still think my research is about breakfast cereal, and have been proud and supportive every step of the way. To Dad for his enthusiastic proof reading and the all-important corrections on corrections, and to Mum for showing off my first paper so proudly.

Thank you to all fellow postgrads, and inhabitants of G26 past and present. Keelia Scott for coffee, cupcakes and CLOPx2, and Maggie Lieu for never failing to surprise me. Richard Pearson for countless Spar trips, pub trips, and teaching me to laugh when things go wrong, and Miranda Bradshaw for keeping me sane with jokes that quite frankly have no right being funny. And Melissa Gillone, for cards and road trips, post offices and December sunburn, overly dramatic TV, and all the great things in between. Sláinte!

Thank you Felicia Ziparo for always being there, and for perfecting with me the fine art of strolling. Mark Levy, Matt Hunt and Jack Gartlan for very welcome distractions, and Steve Gillman, Jonny Clarke and Nathan Adams for giving me a new view of teamwork.

Graham Smith for the opportunity to do some really cool science. Bill Chaplin for continued support and belief in me. Sean McGee for infectious enthusiasm, unrivalled generosity, and questionable ability to fit a line to data. Sara Lawrence and Jo Cox for so much behind the scenes help, and David Stops for somehow keeping me from computing crisis.

Thank you to Team 33 for making Birmingham feel like home way back when, and Nicola Woodcock for keeping it that way. The Best Third Team Ever for more than living up to their name. YouTube and imgur for doing their best to derail my work, and coffee for creeping unnoticed into my life to keep me on track.

And finally, thank you Sean, for believing in me and helping me believe in myself. Thank you for sharing the highs, picking me up when I stumbled, and for reminding me that digging is best. MLM

Statement of originality

I declare that the work submitted in this thesis is my own. Chapters 1 and 2 are fully referenced introductory chapters, while Chapters 3, 4 and 5 are each taken in their entirety from papers of which I am first author. As such all work is original, clearly referenced, or mentioned specifically below.

- Chapter 3 is taken from Mulroy et al. (2014): "LoCuSS: The Near-Infrared Luminosity and Weak-Lensing Mass Scaling Relation of Galaxy Clusters", and is all my own work.

- Chapter 4 is taken from Mulroy et al. submitted: "Galaxy Cluster Luminosities and Colours, and their Dependence on Mass and Merger State", and includes $c_{\rm SB}$ measurements and values calculated from the Millennium simulation by co-authors.

- Chapter 5 is taken from Mulroy et al. in preparation: "LoCuSS: Galaxy Cluster Scaling Relations", and involves the use of a Bayesian code developed by a co-author and X-ray & SZ parameters calculated by co-authors.

Contents

List of Figures

List of Tables

Chapter 1

Cosmology

1.1 Geometry of the Universe

In late 1915, Albert Einstein published a series of papers on his General Theory of Relativity, describing the geometry and the evolution of the Universe. In the century since, observational confirmation of its predictions have proved it to be an accurate description of the Universe, from the observation of gravitational lensing in 1919 (Dyson et al., 1920), to the first direct detection of gravitational waves in 2015 (Abbott et al., 2016).

Einstein's theory allows us to understand the Universe on cosmological scales. A foundation of cosmology is the Cosmological Principle: On large enough scales the Universe is the same everywhere. This requires the Universe to be homogeneous - the same in all positions; and isotropic - the same in all directions. There is no special place in the Universe.

Hubble's Law, $v = Hr$, was constrained using observations of galaxies, and demonstrates that the recession velocity of objects v scales with their distance r by Hubble's Parameter H (Hubble, 1929). Combined with the Cosmological Principle, this tells us that the Universe must be expanding. To deal with that mathematically it is useful to introduce a coordinate system that expands with it. These are comoving coordinates, $r = a(t)x$, where x is the physical coordinate system, and $a(t)$ the all-important scale factor.

70 years after Hubble, two groups (Riess et al., 1998; Perlmutter et al., 1999) tested his law to larger distances using type Ia supernovae. As the absolute magnitude of these supernovae is known, their apparent magnitude provides a measure of their distance. It was found that objects at large distances had lower recession velocities than expected, so the expansion of the Universe must be accelerating. This is driven by Dark Energy, the nature of which remains a mystery.

1.1.1 Friedmann Equation

Employing the Cosmological Principle, we can view the Universe as a uniform medium with density ρ, within which the behaviour of any two test particles represents the behaviour of the whole medium. Conservation of energy tells us that the total energy E, the sum of the kinetic energy K and the potential energy U, must be constant:

$$\begin{aligned} E &= K + U && (1.1)\\ &= \frac{1}{2}m\dot{r}^2 + -\frac{GMm}{r} && (1.2)\\ &= \frac{1}{2}m\dot{r}^2 - \frac{G4\pi r^3\rho m}{3r} && (1.3)\\ &= \frac{1}{2}m\dot{a}^2x^2 - \frac{G4\pi a^2x^2\rho m}{3}, && (1.4) \end{aligned}$$

where G is the Gravitational Constant, m is the mass of a test particle, r the distance between test particles and M the mass within r.

Rearranging gives what is called the Friedmann Equation:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{k}{a^2}, \tag{1.5}$$

where $k = -\frac{2E}{mx^2}$.

It is important to note that the density ρ represents the density of all the components of the Universe: radiation, matter and dark energy:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}(\rho_r + \rho_m + \rho_\Lambda) + \frac{k}{a^2}. \tag{1.6}$$

This equation demonstrates that the geometry of the Universe is related to the density within it, and introduces us to the curvature, k. Due to our assumption of homogeneity, k must be the same everywhere and is a fundamental property of the Universe.

A universe with $k > 0$ is 'closed', with a density high enough that its gravitational attraction will eventually stop and reverse its expansion. A universe with $k < 0$ is 'open', without sufficient density for the gravitational attraction to stop its expansion. A universe with $k = 0$ is 'flat',

with a critical density that causes the expansion to tend to zero as time tends to infinity. This critical density is defined simply by setting $k = 0$ in the Friedmann Equation and rearranging:

$$\rho_c \equiv \frac{3H^2}{8\pi G}. \tag{1.7}$$

1.1.2 Fluid Equation

To understand the behaviour of a material in the Universe, we start with the First Law of Thermodynamics and derive the Fluid Equation. The First Law of Thermodynamics is the principle of conservation of energy applied to thermodynamics, relating the time derivatives of the total energy E, volume V, and entropy S of a system:

$$\dot{E} + p\dot{V} = T\dot{S}, \tag{1.8}$$

where T is the temperature and p the pressure. In a reversible expansion there is no change in entropy, so $\dot{S} = 0$. Scaling c=1 and x=1 allows us to derive the following expressions:

$$E = m = \frac{4\pi r^3}{3}\rho = \frac{4\pi a^3}{3}\rho, \qquad \dot{E} = \frac{4\pi}{3}(3a^2\dot{a}\rho + a^3\dot{\rho}), \tag{1.9}$$

$$V = \frac{4\pi r^3}{3} = \frac{4\pi a^3}{3}, \qquad \dot{V} = 4\pi\dot{a}a^2. \tag{1.10}$$

Combining these equations and rearranging gives us the Fluid Equation:

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \tag{1.11}$$

We can see that the evolution of the density ρ is related to the evolution of the scale factor. The factor in the brackets relating the two is comprised of a term describing the dilution of the density due to expansion and a term describing the loss of energy due to work done by the pressure during expansion.

1.1.3 Equation of State

The Equation of State of a material is the relation between its pressure p and density ρ, and can be defined as

$$p = p(\rho) \equiv w\rho, \tag{1.12}$$

where w depends on the material being considered. Substituting this into the Fluid Equation and rearranging leads to

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + w\rho) = 0, \tag{1.13}$$

$$\frac{d\rho}{\rho} = -3(1+w)\frac{da}{a}, \tag{1.14}$$

which can be integrated to give

$$\rho \propto a^{-3(1+w)}. \tag{1.15}$$

This relation applies to all components of the Universe: radiation, matter and dark energy. Their individual expressions vary due to different Equations of State, and so different values of w.

- For radiation, i.e. relativistic particles, $w = 1/3$, which can be derived from the Equipartition Theorem or conceptually seen by considering that the momentum vectors of individual photons are isotropically distributed in three-dimensional space. This leads to $\rho_r \propto a^{-4}$, where three factors are explained by the expansion of space, and one by the stretching of the wavelength itself.

- For matter, i.e. non-relativistic particles, there is no pressure and so $w = 0$. This leads to $\rho_m \propto a^{-3}$, explained by the expansion of three-dimensional space.

- For dark energy we assume the leading interpretation, that is it a Cosmological Constant Λ (Planck Collaboration et al., 2016b). This was first introduced by Einstein as a term in his equations to keep the Universe static, before it was shown that the Universe was expanding and he famously called it his 'greatest blunder'. Within this interpretation, $w = -1$ and $\rho_\Lambda \propto a^0$. This counter intuitive result tells us that dark energy has negative effective pressure, and it remains constant despite the expansion of space. It is sometimes

thought of as a 'vacuum energy' or a 'zero-point energy', but there is no accepted physical interpretation (Weinberg, 1989; Frieman et al., 2008). An alternative to the Cosmological Constant explanation is that of Quintessence, which we will not explore here, but that postulates that dark energy is a scalar field and does in fact vary slowly with time (Ratra & Peebles, 1988; Caldwell et al., 1998).

For each component we can define corresponding density parameters Ω, and an equivalent for curvature:

$$\Omega = \frac{\rho}{\rho_c} = \frac{8\pi G\rho}{3H^2}, \qquad \Omega_k = \frac{-k}{H^2}. \tag{1.16}$$

Dividing the Friedmann Equation by H^2 explicitly demonstrates that the fractional densities of all the components must sum to one: $1 = \Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda$.

Dividing the Friedmann Equation by H_0^2 gives

$$\frac{H^2}{H_0^2} = \frac{\Omega_{r,0}}{a^4} + \frac{\Omega_{m,0}}{a^3} + \frac{\Omega_{k,0}}{a^2} + \Omega_{\Lambda,0}, \tag{1.17}$$

where a subscript 0 means the current values, and the equation shows which components dominated at different stages of the Universe's expansion. This evolution can be seen in Figure 1.1. For example, we can see that the very early Universe was radiation dominated, and that dark energy starts to dominate as the Universe expands. Current values are $\Omega_{r,0} \simeq 10^{-6}$, $\Omega_{m,0} \simeq 0.3$, $\Omega_{k,0} \simeq 0.0$ and $\Omega_{\Lambda,0} \simeq 0.7$ (Planck Collaboration et al., 2016b).

Taking the equation for redshift and differentiating;

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{v}{c}, \qquad \frac{d\lambda}{\lambda} = \frac{dv}{c} = \frac{\dot{a}}{a}\frac{dr}{c} = \frac{\dot{a}}{a}dt = \frac{da}{a}, \tag{1.18}$$

we see that $\lambda \propto a$, and so

$$1 + z = \frac{\lambda_o}{\lambda_e} = \frac{a_o}{a_e}, \tag{1.19}$$

where λ_e is the emitted wavelength and λ_o the observed wavelength. Substituting $a = \frac{1}{1+z}$ (scaling a such that $a_0 = 1$), we can derive the Evolution Parameter:

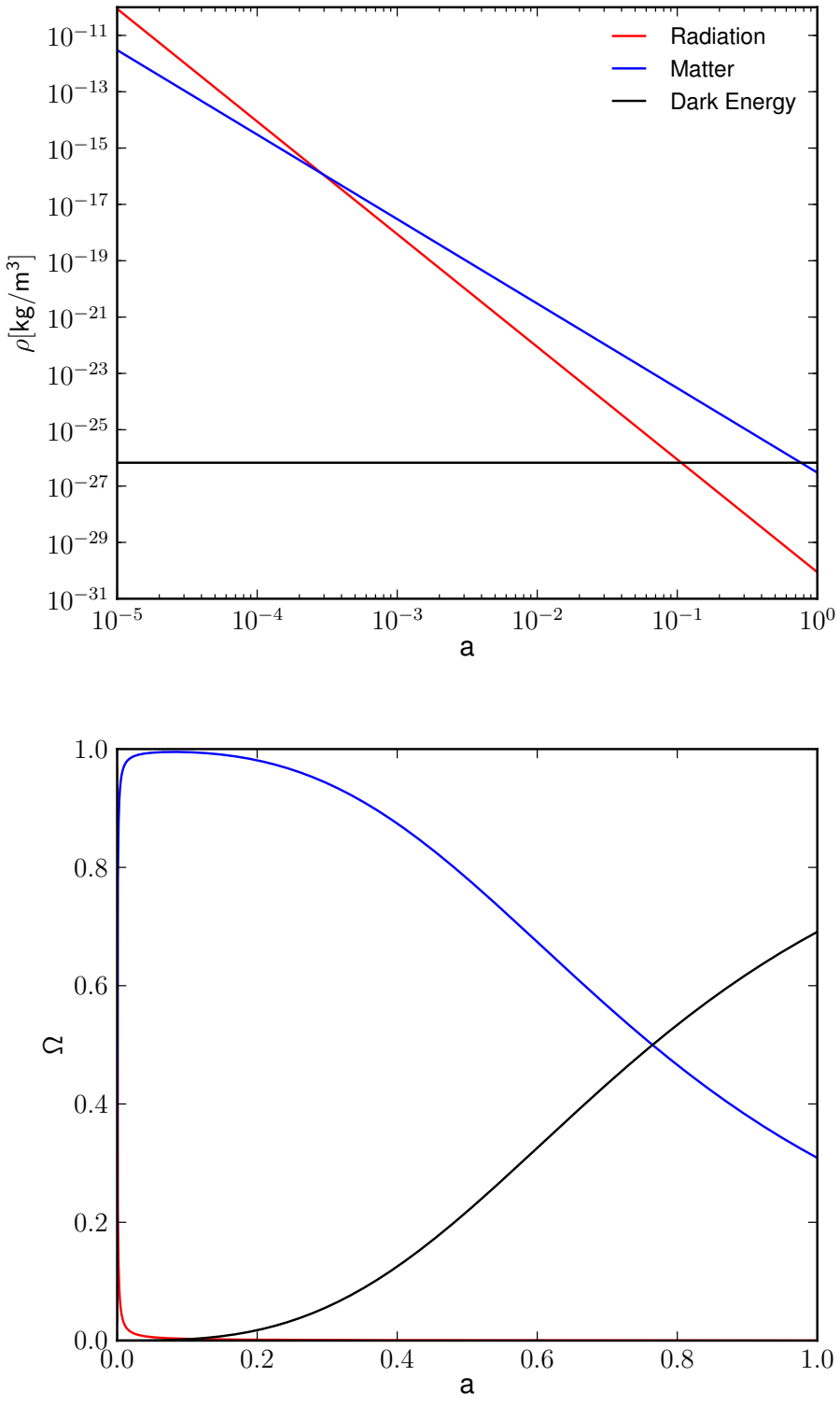

Figure 1.1: The evolution of the density of the three components of the Universe with the scale factor a. Showing radiation (red, $\rho_r \propto a^{-4}$), matter (blue, $\rho_m \propto a^{-3}$) and dark energy (black, $\rho_\Lambda \propto a^0$). Top panel: evolution of density in units of kg/m^3. Lower panel: evolution of density in units of the critical density.

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_{r,0}(1+z)^4 + \Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}}. \tag{1.20}$$

As we will see in Section 2.2.2, this important parameter allows us to compare objects at different redshifts by taking care of the evolution of the critical density:

$$\rho_c(z) = \frac{3H(z)^2}{8\pi G} = \frac{3H_0^2 E(z)^2}{8\pi G} = E(z)^2 \rho_{c,0}. \tag{1.21}$$

1.2 Structure Formation

We have seen above how the Universe and its components evolve on cosmological scales, but the existence of structure tells us that the Cosmological Principle does not hold on smaller scales. In this section we look at how those structures formed in the early Universe, and how they evolved into the complex cosmic web that we observe today.

1.2.1 Growth of Density Fluctuations

In this section we introduce the Spherical Collapse Model, which was first considered by Gunn & Gott (1972). As can be seen clearly in the lower panel of Figure 1.1, the radiation dominated phase of the Universe was very short. We can therefore approximate the early Universe within which structure begins to form as an Einstein de Sitter Universe (EdS, flat and matter dominated). We also assume that this matter is collisionless, which is valid since the majority of matter is in the form of dark matter. Assuming that the seeds of gravitationally bound structures are spherical density fluctuations, δ, in the early Universe gives

$$\delta(r,t) \equiv \frac{\rho(r,t) - \bar{\rho}}{\bar{\rho}}, \tag{1.22}$$

where $\bar{\rho}$ is the average density. This density evolves as an EdS Universe:

$$r_{\rm EdS}(t) = \frac{A}{2}\left(\frac{6t}{B}\right)^{2/3}, \tag{1.23}$$

where $A^3 = GMB^2$ and $r_{\rm EdS}$ is a unit radius where we have set $x = 1$.

The evolution of the overdense sphere decouples from that of the Hubble flow of Equation 1.23 and evolves like a closed universe ($k > 0$), reaching a maximum radius and collapsing. The parametric solutions are:

$$r(t) = A(1 - \cos\theta), \tag{1.24}$$

$$t = B(\theta - \sin\theta), \tag{1.25}$$

and

$$\delta(t) \simeq \frac{3}{20}\left(\frac{6t}{B}\right)^{2/3}. \tag{1.26}$$

Examining Equation 1.24, the maximum possible radius of the sphere is $r_{\rm max} = 2A$, occurring when $\theta = \pi$, $t_{\rm max} = \pi B$ and $\delta_{\rm max} = 1.06$. It then collapses under gravity to $r_{\rm coll} = 0$, occurring when $\theta = 2\pi$, $t_{\rm coll} = 2\pi B$ and $\delta_{\rm coll} = 1.69$. Collapse to $r_{\rm coll} = 0$ is clearly not physical; instead the sphere will virialise, satisfying $U = -2K$. To work out the radius at which this happens is simple: at $r_{\rm max}$ the total energy of the system $E = U + K = U$, while at $r_{\rm vir}$ the total energy of the system $E = U + K = U - \frac{1}{2}U = \frac{1}{2}U$. Conservation of energy tells us that $U(r_{\rm max}) = \frac{1}{2}U(r_{\rm vir})$, and since $U \propto \frac{1}{r}$, $r_{\rm vir} = \frac{1}{2}r_{\rm max}$. This occurs when $\theta = \frac{3}{2}\pi$, $t_{\rm vir} = (\frac{3}{2}\pi + 1)B$ and $\delta_{\rm vir} = 1.58$, giving a cosmic overdensity:

$$\Delta_{\rm vir} = \frac{\rho_{\rm vir}}{\bar{\rho}(t_{\rm vir})} = \left(\frac{r_{\rm EdS}(t_{\rm vir})}{r_{\rm vir}}\right)^3 \simeq 5.5. \tag{1.27}$$

It is more common to allow the background to continue expanding until $t_{\rm coll}$, so the virial overdensity becomes

$$\Delta_{\rm vir} = \frac{\rho_{\rm vir}}{\bar{\rho}(t_{\rm coll})} = \left(\frac{r_{\rm EdS}(t_{\rm coll})}{r_{\rm vir}}\right)^3 = 18\pi^2 \simeq 178. \tag{1.28}$$

From this definition, an object is commonly identified as gravitationally collapsed only when its average density is at least 178 times the background density of the Universe.

1.2.2 Mass Function

In order to relate the density fluctuations in the early Universe to the structure we observe today, we need a formalism for the halo mass function. The mass function is the number density

of halos in a given mass interval, and we derive it below following the pioneering work of Press & Schechter (1974).

Assuming that the smoothed overdensity field $\delta(r,t)$ is a Gaussian random field with variance $\sigma = \langle\delta\rangle$, then the probability that an overdensity has a value above a critical overdensity δ_c is:

$$P(>\delta_c) = \frac{1}{\sqrt{2\pi}\sigma}\int_{\delta_c}^{\infty}\exp\left(-\frac{\nu^2}{2}\right)d\delta = \frac{1}{2}\mathrm{erfc}\left(\frac{\nu_c}{\sqrt{2}}\right), \tag{1.29}$$

where $\nu \equiv \delta/\sigma$ is the overdensity in units of the variance, and $\nu_c \equiv \delta_c/\sigma$.

The Press & Schechter ansatz is that this probability, $P(>\delta_c)$, is equal to the fraction of mass elements residing in halos with mass greater than M, $F(>M)$. However a problem arises when $\sigma \to \infty$, leading to $P(>\delta_c) \to 1/2$. If $F(>M) = P(>\delta_c)$, this suggests that only half the matter in the Universe resides in a collapsed object. The simplest physical explanation is that the linear theory only considers positive fluctuations, and so only half the matter. Press & Schechter corrected for this by multiplying by an ad hoc factor of two, leading to $F(>M) = 2P(>\delta_c)$.

The interesting quantity is $dn = n(M)dM$, the number density of objects within a mass range $[M, M+dM]$. $F(M)dM$ is the fraction of objects in the same mass range, and when multiplied by ρ is the total mass within the range. It therefore follows that:

$$\begin{aligned}
n(M)dM &= \frac{\rho}{M}F(M)dM && (1.30)\\
&= \frac{\rho}{M}\frac{\partial F(>M)}{\partial M}dM && (1.31)\\
&= \frac{2\rho}{M}\frac{\partial P(>\delta_c)}{\partial M}dM && (1.32)\\
&= \frac{2\rho}{M}\frac{\partial P(>\delta_c)}{\partial \sigma}\left|\frac{d\sigma}{dM}\right|dM && (1.33)\\
&= \frac{2\rho}{M^2}\frac{\partial P(>\delta_c)}{\partial \sigma}\left|\frac{d\ln\sigma}{d\ln M}\right|dM && (1.34)\\
&= \frac{2\rho}{M^2}\frac{\nu_c}{\sqrt{2\pi}}\exp\left(-\frac{\nu_c^2}{2}\right)\left|\frac{d\ln\sigma}{d\ln M}\right|dM && (1.35)\\
&= \frac{\rho}{M^2}\sqrt{\frac{2}{\pi}}\nu_c\exp\left(-\frac{\nu_c^2}{2}\right)\left|\frac{d\ln\sigma}{d\ln M}\right|dM && (1.36)\\
n(M)dM &= \frac{\rho}{M^2}f_{PS}(\nu_c)\left|\frac{d\ln\sigma}{d\ln M}\right|dM, && (1.37)
\end{aligned}$$

where

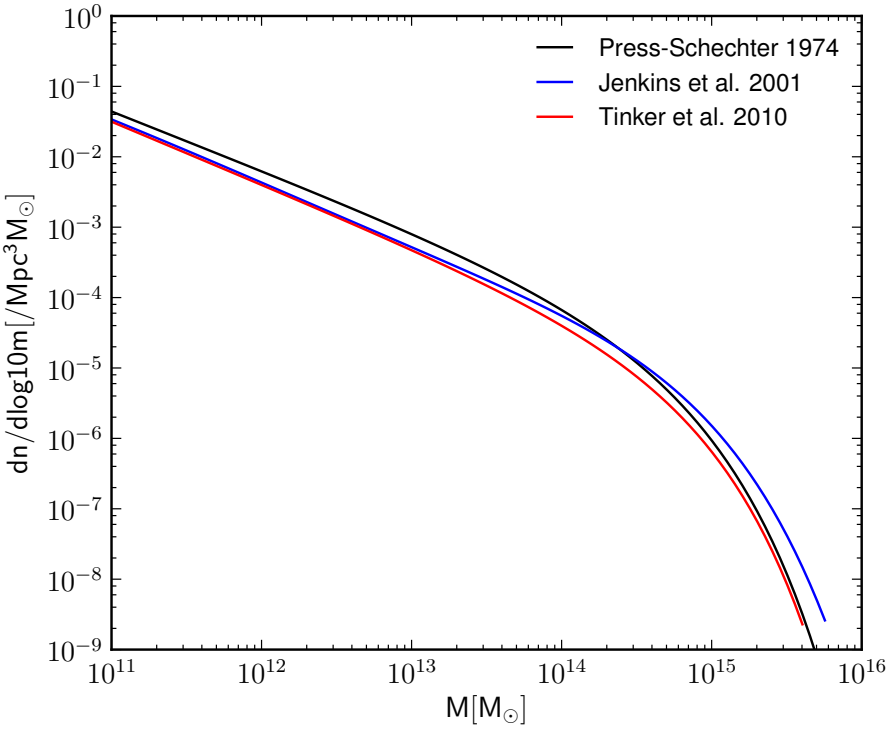

Figure 1.2: The halo mass functions of Press & Schechter (1974), Jenkins et al. (2001) and Tinker et al. (2010).

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}}\nu e^{-\nu^2/2} \tag{1.38}$$

is the multiplicity function, the fraction of mass in a unit range of $\ln \nu$.

While this approach is rather simple, it does a remarkably good job of reproducing the observed halo distribution. Numerical simulations have since been used to derive the mass function, based on gravity and following the evolution of fluctuations through time (e.g. Jenkins et al., 2001; Tinker et al., 2010). As shown in Figure 1.2, these simulations agree well with each other and with the Press & Schechter formalism.

The expressions for the halo mass function are dependent on the underlying cosmology. Therefore by varying the cosmological parameters and comparing to observations, we can constrain these parameters. As can be seen in Figure 1.2, and from the form of the mass function in Equation 1.37, the mass function at the high mass end is exponentially decaying, and so this region has the most constraining power. This part of the mass function is populated by galaxy clusters - the largest gravitationally bound objects in the Universe. It is therefore important for

cosmology to make highly accurate and precise mass measurements of these objects (e.g. Allen et al., 2011), which we will discuss in Chapter 2.

Chapter 2

Galaxy Clusters

In the 1780s, overdensities of galaxies on the sky were identified in the Messier Catalogue (Messier, 1781). We now know these as galaxy groups and clusters - gravitationally bound collections of tens to thousand of galaxies. While there is no definitive distinction between groups and clusters, groups are the smaller of the two, containing typically only a few ($\lesssim 10$) galaxies. Originally identified by their galaxies in the optical, groups and clusters also contain hot gas in the intracluster medium (ICM) which emits in the X-ray. In 1933, Zwicky found that the orbital velocities of galaxies in the Coma Cluster were much greater than could be supported by the gravitational potential of the visible mass in the cluster. This 'missing mass' is Dark Matter, which we now know accounts for the majority of matter in galaxy groups, galaxy clusters, and the Universe as a whole.

Galaxy clusters are so large that their matter content is thought to have the same composition as that of the Universe: $\sim$ 85% dark matter and $\sim$ 15% baryonic matter. Their baryonic content is further broken down into $\sim$ 10% gas in the ICM, and $\sim$ 5% stellar material mostly in the galaxies (e.g. Gonzalez et al., 2013). A typical galaxy cluster mass is $\sim 10^{15} M_{\odot}$, which is dominated by dark matter, with the stellar mass making up $\sim 10^{13} M_{\odot}$ and the gas mass $\sim 10^{14} M_{\odot}$. This gas is typically at a temperature of $\sim$ 7 keV, and emits an X-ray luminosity of $\sim 10^{45}$erg/s.

2.1 Mass Measurements

As galaxy clusters populate the high mass end of the mass function, their mass measurements are important for constraining cosmology. There are a variety of methods to make these measurements, each with their own advantages and disadvantages. We discuss the three main direct mass measurement methods below.

2.1.1 Dynamical

Dynamical mass measurements involve using the galaxies within clusters as tracers of the gravitational potential of the system. Cluster galaxies reside in a virialised system, satisfying the the Virial Theorem: the relation between the time average potential energy $\langle U \rangle$ and the time average kinetic energy $\langle K \rangle$, given by

$$\langle U \rangle = -2\langle K \rangle. \tag{2.1}$$

These energies can be expressed as:

$$\langle U \rangle = -\alpha \frac{GMM_{\rm gal}}{r}, \qquad \langle K \rangle = \sum_{\rm gal} \frac{1}{2} m_{\rm gal} v^2 = \frac{1}{2} M_{\rm gal} \langle v^2 \rangle = \frac{3}{2} M_{\rm gal} \langle v_{1D}^2 \rangle = \frac{3}{2} M_{\rm gal} \sigma^2, \tag{2.2}$$

where M is the total mass of the cluster, $M_{\rm gal}$ the total mass in galaxies, and $m_{\rm gal}$ the mass of an individual galaxy which we assume to be the same for all galaxies. v is the galaxy velocity, $v_{\rm 1D}$ the galaxy velocity along one axis, and the velocity dispersion σ is equal to the standard deviation of the galaxy velocities along the line of sight. The value of α depends on the density profile of the cluster density, and for example equals 3/5 under the assumption of uniform density, and 3 under the assumption of $\rho \propto r^{-2}$. Combining these equations leads to the relation between velocity dispersion and mass:

$$M = \frac{3r\sigma^2}{\alpha G}. \tag{2.3}$$

The observations required for this mass measurement are spectra of ideally all cluster members, or enough that the population of velocities is sampled well enough to derive an accurate measure of the velocity dispersion. At the low redshifts of the clusters studied in this thesis ($0.15 < z < 0.30$), the time requirements are moderate; ~ 250 galaxies can be targeted in ~ 1 hour with a single pointing on a 6.5m telescope (Haines et al., 2013). However to observe the ~ 50 galaxies required takes ~ 15 hours with an $8-10$m telescope at higher redshifts ($z \sim 1$) (Muzzin et al., 2012). This measurement could suffer from biases if the distribution of the velocity of cluster galaxies is not well fit by a Gaussian.

2.1.2 Hydrostatic

Hydrostatic mass measurements assume that the gas within a cluster is in hydrostatic equilibrium - that the gravitational pressure acting to collapse the cluster is exactly equalled by the gas pressure driving it apart. Under hydrostatic equilibrium the relation between the gas pressure P and the density ρ is

$$\frac{dP}{dr} = -\frac{GM(r)}{r^2}\rho. \tag{2.4}$$

Assuming the ideal gas law $P = \frac{k_B}{\mu m_H}\rho T$ and rearranging, this gives us an expression for the mass of the cluster:

$$M(r) = -\frac{kT}{\mu m_H G}\left(\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln R}\right)r, \tag{2.5}$$

where k_B is the Boltzmann Constant, m_H the mass of a hydrogen atom and μ the mean molecular weight. X-ray emission is dominated by bremsstrahlung emission, allowing us to measure the temperature from the spatially resolved spectra of X-ray observations, and the density from the flux. At low redshifts these measurements require ~ 3 hours of observations on an X-ray satellite per cluster (Cavagnolo et al., 2009), while at higher redshifts the time required is closer to ~ 30 hours per cluster (Bartalucci et al., 2017). This measurement could suffer from biases if the assumption of hydrostatic equilibrium and spherical symmetry do not hold. Recent work on the Perseus Cluster found little turbulence in the hot gas in the central regions of the cluster, suggesting that the assumption of hydrostatic equilibrium is valid (Hitomi Collaboration et al., 2016).

2.1.3 Gravitational Weak-Lensing

Gravitational lensing involves studying the effect of the cluster's gravitational potential on the distribution of light from distant galaxies. In the weak-lensing regime this distortion is small, and the mass measurement requires a large number of background galaxies (e.g. Bartelmann & Schneider, 2001).

The path of the light is distorted by the gravitational field, with a deflection angle

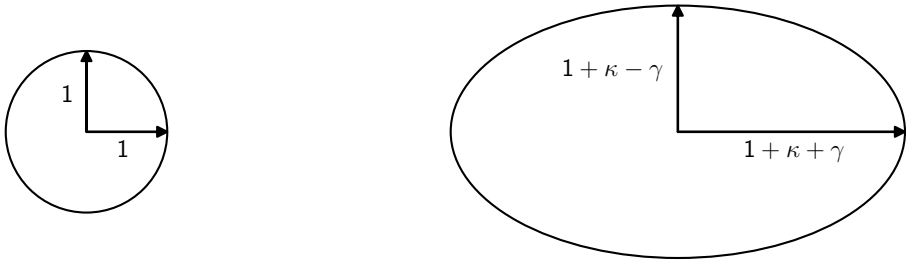

Figure 2.1: The effect on a single galaxy of the convergence κ and the shear γ. The circle [left] shows the original unlensed source, while the ellipse [right] shows the source after it has been distorted by a gravitational potential.

$$\alpha \propto \frac{M}{R}, \tag{2.6}$$

where M is the deflecting mass, and R the impact parameter between the light path and that mass. The Jacobian matrix, which relates the lensed and unlensed plane is

$$A = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}, \tag{2.7}$$

where ϕ takes care of the angle of deformation with respect to the x-axis. The convergence κ is an isotropic deformation - rescaling the source the same in all directions, and the shear γ is an anisotropic deformation - stretching the shape along a preferred direction. The result, as shown in Figure 2.1, is a combination of the two from which we are unable to measure the separate distortions. The best we can measure is the reduced shear $g = \gamma/(1-\kappa)$, although in the weak-lensing regime $\kappa << 1$ so $g \simeq \gamma$. Any intrinsic ellipticity in the galaxies, $\epsilon_{\rm int}$, is assumed on average to sum to zero, and since weak-lensing is a statistical measurement we can consider the measured ellipticities of the galaxies, $\epsilon_{\rm obs}$, to be a measure of the shear:

$$\epsilon_{\rm obs} = g + \epsilon_{\rm int} \simeq g \simeq \gamma. \tag{2.8}$$

The light is preferentially distorted tangentially to the cluster centre as shown in Figure 2.2, and so the gravitational potential can be reconstructed from these measurements of the

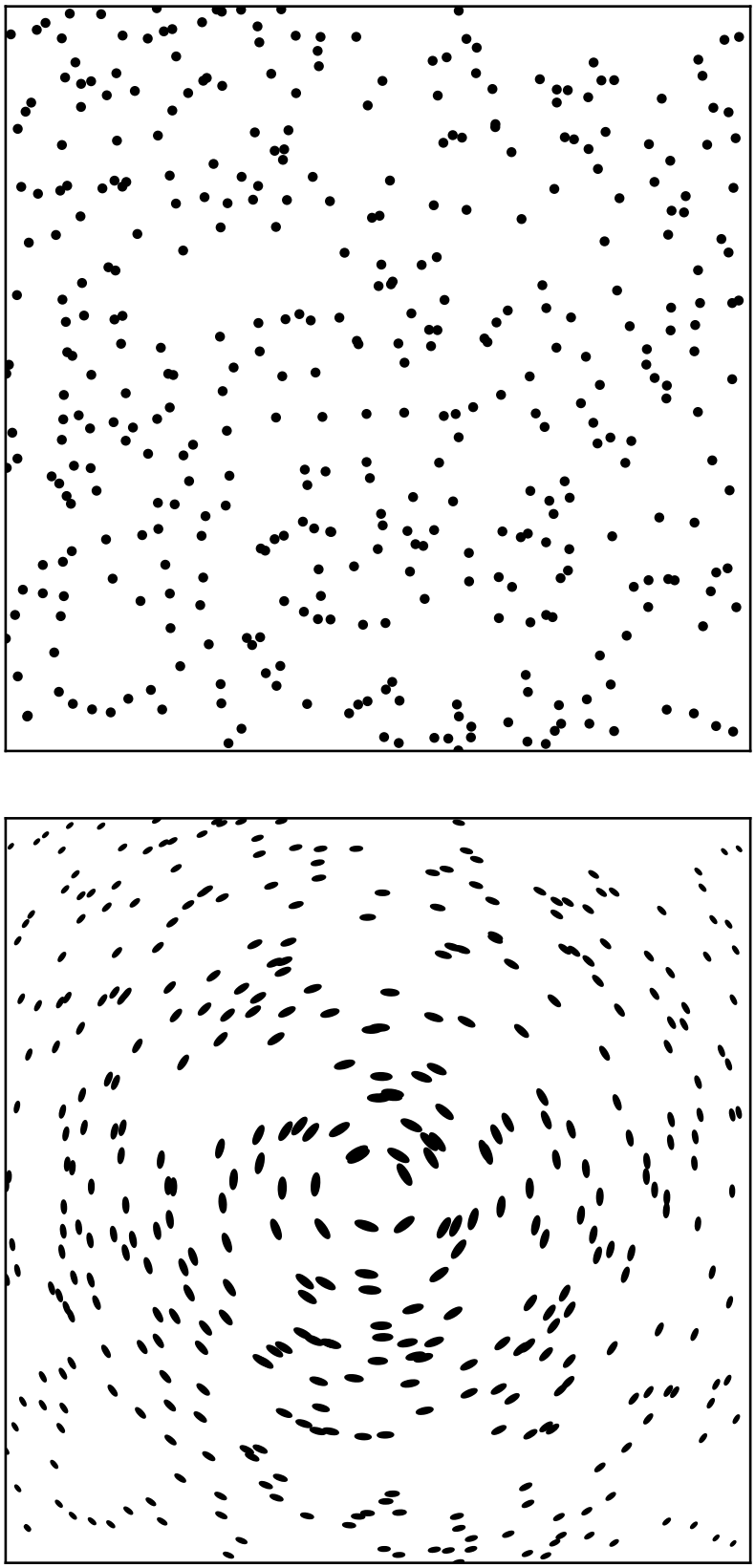

Figure 2.2: Schematic of the distortion of light from background galaxies caused by gravitational weak-lensing by a massive object. Top: light from background galaxies in the absence of a massive object. Lower: light from background galaxies having been distorted by an otherwise unseen massive object in the centre of the plot.

shear, and a model of the mass distribution can be fitted. The best fitting model predicted by simulations is the NFW model (Navarro, Frenk & White, 1997) which has a density of

$$\frac{\rho}{\rho_s} = \frac{1}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}, \tag{2.9}$$

where both ρ_s and the scale radius r_s vary cluster to cluster. The density is $\propto r^{-1}$ towards the centre, $\propto r^{-2}$ like an isothermal sphere at $r = r_s$, and $\propto r^{-3}$ at the outskirts of the cluster

This measurement requires deep enough observations that the shape of the background galaxies can be measured, and a background density large enough to provide the required statistics without having to stack clusters. The time requirement for a low redshift cluster is ~ 1 hour on an $8-10$m telescope (Okabe et al., 2010a), with higher redshifts needing several orbits on the Hubble Space Telescope (Jee et al., 2011).

As a geometric calculation, this is the only mass measurement which does not rely on assumptions about the dynamical state of the cluster, only that it is well fitted by an NFW profile. However, it is sensitive to structure along the line of sight and projection effects. Simulations show that weak-lensing masses are on average unbiased with respect to true mass, with a scatter of $\sigma_{\ln M_{\rm WL}|M_{\rm True}} = 0.2$ (Oguri & Hamana, 2011; Becker & Kravtsov, 2011; Bahé et al., 2012).

2.2 Scaling Relations

While the mass measurement methods above are all very different, requiring a variety of observations and making different assumptions, they all require deep observations and extensive analysis. This is feasible for relatively small samples of galaxy clusters, but not for the large statistical samples required to constrain the mass function and probe cosmology. This motivates an interest in scaling relations between mass and 'mass proxies'. These proxies are cluster observables that do not require such deep observations or such extensive analysis, and can be measured for large numbers of clusters. If we can relate a mass proxy to mass, we can use measurements of this observable to constrain cosmology. An ideal mass proxy is one with minimal intrinsic scatter with mass, so that the mass can be accurately determined from the proxy. These scaling relations are calibrated on a relatively small number of clusters, for which accurate measurements of both the mass and the mass proxy are available.

2.2.1 Linear Regression

The scaling relation between cluster observable and mass is normally assumed to be a power law. This is a straight line in logarithmic space, and so a scaling relation is parameterised by performing linear regression on the logarithmic values:

$$Y = \alpha X^{\beta}, \tag{2.10}$$

$$y = bx + a, \tag{2.11}$$

where the independent variable $x = \ln X$ and the dependent variable $y = \ln Y$, and with intercept a ($\ln \alpha$), slope b (β), and intrinsic scatter $\sigma_{\ln Y|X}$.

There are many factors for a linear regression method to consider, including but not limited to intrinsic scatter, observational errors, and selection effects.

Many approaches to linear regression involve minimising χ^2; finding the line that minimises the sum of the squares of the distances between the line and the observed data points:

$$\chi^2 = \sum_i \frac{(y_i - (bx_i + a))^2}{w_i^2}, \tag{2.12}$$

where w_i^2 is the weighting that varies between methods. Ordinary Least Squares uses equal weighting for each data point. The standard Weighted Least Squares incorporates measurement error in the dependent variable, using $w_i^2 = \sigma_{\rm int}^2 + \sigma_y^2$, while the method of FITEXY (Press et al., 1992; Tremaine et al., 2002) also incorporates errors in the independent variable, using $w_i^2 = \sigma_{\rm int}^2 + \sigma_y^2 + b^2\sigma_x^2$. The method of BCES (Bivariate Correlated Errors and Intrinsic Scatter, Akritas & Bershady, 1996) handles errors in a more complex way, including accounting for covariance between errors. This BCES method is strictly BCES(Y|X), which minimises residuals along the dependent axis. There is also BCES(X|Y) which minimises residuals along the independent axis, a bisector method which bisects the two, and an orthogonal method which minimises orthogonal residuals.

Another approach to linear regression is the Bayesian approach, such as Kelly (2007) and Mantz (2016). These models are more complex, and as such are able to take into account more properties of a real-world dataset. For example, they can handle non-detections, and model

the distribution of the independent variable as a mixture of Gaussians. In some cases it is also possible to correct for selection effects.

It is easy to see the effect on the linear regression parameters of selecting a sample as those with a dependent variable above some selection limit. As shown in the top panel of Figure 2.3, the selection produces a step function in the dependent variable, above which the sample is complete and below which the sample is absolutely incomplete. A naive fit to this data will underestimate the slope of the relation. It is less easy to see the effect of selecting on a third parameter that has some covariance with the dependent variable. As shown in the lower panels of Figure 2.3, the selection no longer produces a step function in the dependent variable, but instead a 'fuzzy' region within which the sample completeness declines. A naive fit will underestimate the slope if the covariance is positive, and overestimate the slope if the covariance is negative. To correct for this it is necessary to perform a multivariate analysis, involving both the selection variable and the dependent variable, and of course the independent variable, as we will explore in Section 5.3.1.

2.2.2 Self-Similar Predictions

Galaxy clusters are rich physics laboratories, within which many astrophysical processes occur. Different components and processes are observable at different wavelengths, and so there are many different cluster observables. By considering self-similarity (Kaiser, 1986), their scaling relations with mass can be predicted as we will see in this section.

Self-similarity is based on the assumption that on the scale of galaxy clusters the dominant force is gravity, which is scale invariant. This means that clusters are simply scaled versions of each other, with their properties determined only by their mass and redshift. The mass determines the strength of the gravitational potential, while the redshift determines the evolution of the critical density as seen in Equation 1.21: $\rho_c(z) = E(z)^2\rho_{c,0}$. This leads to

$$M = \frac{4}{3}\pi r^3 \Delta \rho_c(z) \tag{2.13}$$

$$= \frac{4}{3}\pi r^3 \Delta \rho_{c,0} E(z)^2 \tag{2.14}$$

$$\propto E(z)^2 r^3, \tag{2.15}$$

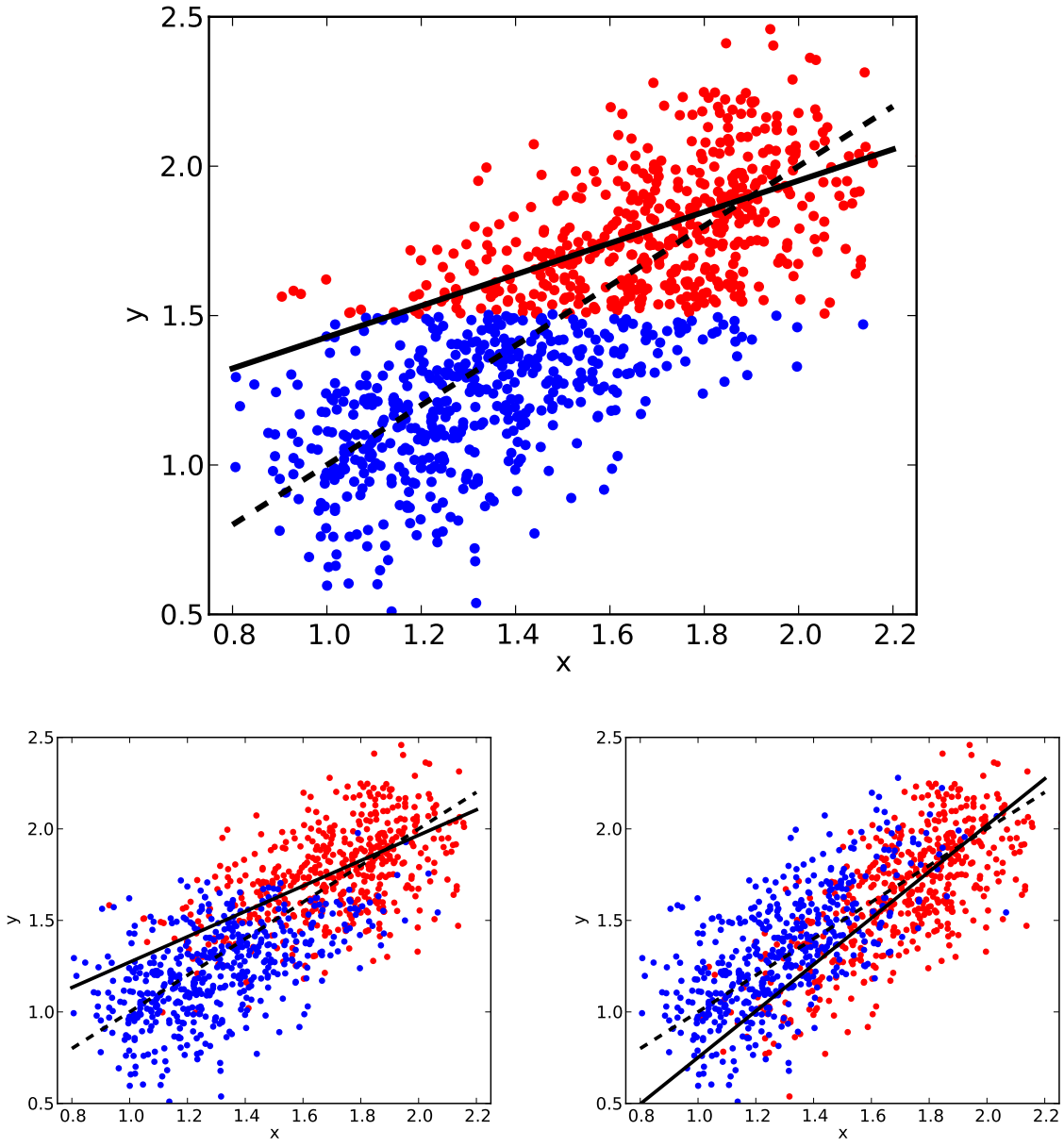

Figure 2.3: A demonstration of how sample selection affects linear regression. In each plot the dashed line is the true scaling relation, and the solid line is the scaling relation fit to the red points using BCES(Y|X), which in all cases is clearly biased. Top: The sample is selected on the dependent variable y and a step function is clearly seen. The naive fit underestimates the slope. Lower left: The sample is selected on a third parameter z with positive covariance with y at fixed independent variable x. Lower right: The sample is selected on a third parameter z with negative covariance with y at fixed independent variable x. The selection is much less clear in these lower two plots, and the scaling relation less biased but still not correct. The naive fit underestimates the slope when the covariance is positive, and overestimates the slope when the covariance is negative.

where the overdensity Δ is the average density inside the sphere in units of the critical density.

Velocity Dispersion

In Section 2.1.1 we saw that galaxy clusters satisfy the Virial Equation $\langle U \rangle = -2\langle K \rangle$. For galaxies:

$$\langle U \rangle = \alpha \frac{GMM_{\rm gal}}{r}, \qquad \langle K \rangle = \sum_{\rm gal} \frac{1}{2} m_{\rm gal} v^2 = \frac{1}{2} M_{\rm gal} \langle v^2 \rangle = \frac{3}{2} M_{\rm gal} \langle v_{\rm 1D}^2 \rangle = \frac{3}{2} M_{\rm gal} \sigma^2, \tag{2.16}$$

which gives our first self-similar scaling relation, between velocity dispersion and cluster mass:

$$\sigma \propto (ME(z))^{1/3}. \tag{2.17}$$

X-ray Temperature

The Virial Equation also applies to the gas in the system, where the total kinetic energy can be written in terms of the average kinetic energy of the ICM particles, i.e. the cluster X-ray temperature, T_X:

$$\langle U \rangle = \alpha \frac{GMM_{\rm gas}}{r}, \qquad \langle K \rangle = \sum_{\rm gas} m_e k_B T_X = M_{\rm gas} k_B T_X, \tag{2.18}$$

where k_B is the Boltzmann Constant and m_e the mass of an electron. Combined with the Virial Equation, this leads to the self-similar relation between cluster temperature and mass:

$$T_X \propto (ME(z))^{2/3}. \tag{2.19}$$

X-ray Luminosity

The X-ray emission from the ICM is dominated by thermal bremsstrahlung emission, in which a free electron is deflected from its path by the field of an ion in the ICM and emits a photon. The resulting luminosity scales as

$$L_X \propto \rho_{\rm gas}^2 r^3 \Lambda(T_X), \tag{2.20}$$

where there are two factors of the gas density $\rho_{\rm gas}$ ($\propto E(z)^2$) because it is a two-body interaction, and $\Lambda(T_X)$ is the cooling function (Boehringer & Hensler, 1989; Sutherland & Dopita, 1993). In the soft-band ($\sim 0.1-2.4$keV), the cooling function is independent of T_X, and so the self-similar relation between soft-band luminosity and mass is:

$$\frac{L_X}{E(z)} \propto ME(z). \tag{2.21}$$

The cooling function across the full energy range used for bolometric X-ray luminosity scales with $T_X^{1/2}$, and so the self-similar relation between bolometric luminosity and mass is:

$$\frac{L_X}{E(z)} \propto (ME(z))^{4/3}. \tag{2.22}$$

Gas Mass

The simple assumption of a constant gas fraction, $F_{\rm gas}$, gives us the relation between the gas mass $M_{\rm gas}$ and total mass:

$$M_{\rm gas} = F_{\rm gas}M \propto M. \tag{2.23}$$

Thermal Energy of the ICM

A final X-ray parameter $Y_X = M_{\rm gas}T_X$ was proposed by Kravtsov et al. (2006), who found it to be a low scatter mass proxy after observing an anti-correlation between $M_{\rm gas}$ and T_X at fixed M in simulations. As the product of the two it is a measure of the total thermal energy of the ICM, and its self-similar relation with mass is

$$Y_X E(z) \propto (ME(z))^{5/3}. \tag{2.24}$$

The same relation holds for Y_{SZ}: the integrated Sunyaev-Zel'dovich (SZ) parameter. Again a measure of the thermal energy of the ICM, this parameter is the result of the inverse Compton scattering of cosmic microwave background (CMB) photons by hot electrons in the ICM. These interactions boost the photon energy by $\sim k_BT/m_ec^2$, increasing the signal at $\gtrsim 218$ GHz,

corresponding to the location of the peak of the CMB spectrum after interaction with the ICM (Carlstrom et al., 2002).

Stellar Mass

As with the gas fraction, the assumption of a constant stellar fraction, $F_\star$, gives us the relation between the stellar mass $M_\star$ and total mass:

$$M_\star = F_\star M \propto M. \tag{2.25}$$

Near-Infrared Luminosity

L_K is the total cluster luminosity in the K band (centred on 2.21μm), which is sensitive to old red stars and relatively insensitive to younger blue stars. It is therefore a measure of the stellar mass (Bell & de Jong, 2001), and so its relation with mass is:

$$L_K \propto M. \tag{2.26}$$

Richness

If we assume each galaxy to have the same mass, $m_{\rm gal}$, we can also derive a relation between richness (the number of cluster galaxies), λ, and mass:

$$\lambda \propto \frac{M}{m_{\rm gal}} \propto M. \tag{2.27}$$

2.2.3 Departures from Self-Similarity

Departures from the self-similar scaling relations suggest that the basic underlying assumptions are not a complete description of cluster physics. The self-similar framework does not take into account any baryonic feedback or turbulence, and assumptions used above include that of the Virial Equation, spherical symmetry and that all cluster galaxies have the same mass. Comparing the scaling relation parameters predicted using the self-similar derivation with fits to observed data can provide an insight into how well these assumptions hold.

2.3 Thesis Structure

The work in this thesis is motivated by the importance of scaling relations to galaxy cluster cosmology as introduced in Chapters 1 and 2. In Chapter 3 I investigate the scaling relation between L_K and M_{WL}. In Chapter 4 I extend this work to consider the scaling relations of luminosities at bluer wavelengths, and investigate trends between cluster colour and indicators of the level of disturbance in the clusters. Finally, in Chapter 5 I use a clearly defined sample of galaxy clusters to fit both scaling relations and covariances between observables across the electromagnetic spectrum, having taken into account the effects of the selection function.

Chapter 3

The Near-Infrared Luminosity and Weak-Lensing Mass Scaling Relation of Galaxy Clusters

This chapter is taken from Mulroy et al. (2014): "LoCuSS: The Near-Infrared Luminosity and Weak-Lensing Mass Scaling Relation of Galaxy Clusters", and is all my own work.

3.1 Introduction

The growth rate and internal structure of galaxy clusters are sensitive to the cosmological model. Clusters are therefore well established cosmological tools that hold much promise for ongoing and imminent cosmological studies, including those that aim to measure the dark energy equation of state (Allen et al., 2011, and references therein). Clusters are tracers of the high mass end of the mass function and so to test cosmological models against observations requires an accurate measurement of the cluster halo mass. As the mass of clusters is dominated by dark matter, this quantity cannot be measured directly and generally requires extensive observations and modelling.

The importance and complexity of cluster mass measurements are among the key motivations for studying scaling relations between mass M and another observable O, or 'mass proxy'. The form of these relations is motivated by predictions from self-similarity (Kaiser, 1986) that they are power laws, parameterised by normalisation a, slope b, and intrinsic scatter $\sigma_{\ln M|O}$. An ideal scaling relation has low intrinsic scatter, while an ideal observable is inexpensive to measure and preferably obtainable from shallow survey data. Also important is a clear understanding of the

relationship between the measured mass and the 'true' mass, and minimal covariance between M and O.

Most scaling relation studies are based on X-ray observations, and thus assume that the intracluster medium is in hydrostatic equilibrium with the cluster potential. Gas mass, $M_{\rm gas}$, and X-ray temperature, T_X, have been shown to be related to the hydrostatic mass of clusters with intrinsic scatter of $\sim 10\%$ and $\sim 15-20\%$ respectively (e.g. Arnaud et al., 2007; Mantz et al., 2010). The pseudo-pressure of the intracluster gas, namely $Y_X = T_X M_{\rm gas}$, was predicted by simulations to be related to hydrostatic mass with an intrinsic scatter as low as 5% (Kravtsov et al., 2006), however observations suggest a figure closer to $\gtrsim 15\%$ (e.g. Arnaud et al., 2007; Mantz et al., 2010; Martino et al., 2014).

Following several early exploratory studies (Smail et al., 1997; Hjorth et al., 1998; Smith et al., 2005; Bardeau et al., 2007), scaling relation studies based on gravitational lensing mass measurements have developed rapidly in the last few years. The advantage of lensing mass measurement is that it makes no assumption about the dynamical and hydrostatic state of the cluster, although it has irreducible scatter of $\sim 20-30\%$ due to projection effects and uncorrelated large-scale structure along the line of sight (e.g. Meneghetti et al., 2010; Becker & Kravtsov, 2011; Bahé et al., 2012; Rasia et al., 2012). Lensing-based results generally agree with X-ray-based studies that $M_{\rm gas}$ is the lowest scatter X-ray mass proxy, with $\sim 10-15\%$ intrinsic scatter (e.g. Okabe et al., 2010b; Mahdavi et al., 2013), with Y_X presenting $\sim 20-25\%$ scatter (e.g. Okabe et al., 2010b; Mahdavi et al., 2013). Recent measurements of the scaling relation between weak-lensing mass and the integrated Compton parameter, Y_{SZ}, find intrinsic scatter of $\sim 10-20\%$ (Marrone et al., 2012; Hoekstra et al., 2012), in broad agreement with Sunyaev-Zel'dovich effect studies that employ hydrostatic mass estimates (e.g. Bonamente et al., 2008; Andersson et al., 2011).

The integrated optical/near-infrared luminosity of the cluster galaxies can also be used as a mass proxy. K band luminosity is a well-known and reliable tracer of the stellar mass in galaxies, as it is sensitive to old stars and relatively insensitive to more recent star formation and dust extinction (Kauffmann & Charlot, 1998). Several studies have investigated near-infrared luminosity, finding that the $M-L_K$ scaling relation has a scatter of $\gtrsim 30\%$ (e.g. Lin et al., 2003, 2004; Ramella et al., 2004; Rines et al., 2004; Muzzin et al., 2007). They have all used

either dynamical or X-ray mass measurements. In contrast, strong- and weak-lensing studies of clusters report that near-infrared luminosity traces the density and structure of clusters to good accuracy (Kneib et al., 2003; Smith et al., 2005; Richard et al., 2010). These results suggest that the relationship between weak-lensing mass and near-infrared luminosity may have a lower scatter than that between X-ray/dynamical mass and near-infrared luminosity.

In this chapter we present a pilot study of the scaling relation between weak-lensing mass and K band luminosity for a sample of 17 clusters at $0.15 < z < 0.3$. We summarise the gravitational weak-lensing masses and calculate the K band luminosities in Section 3.2. The results are presented in Section 3.3, compared with other published results in Section 3.4, and our findings summarised in Section 3.5. All photometric measurements are relative to Vega, and we assume $\Omega_{\mathrm{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In this cosmology, at the average cluster redshift, $\langle z \rangle = 0.23$, 1 arcsec corresponds to a projected physical scale of 3.67 kpc.

3.2 Data and Analysis

3.2.1 Sample

We study a sample of 17 X-ray luminous clusters at $0.15 < z < 0.3$ (Table 3.1) that have featured in a series of papers from the Local Cluster Substructure Survey (LoCuSS[1]). They are those with weak-lensing masses published in Okabe et al. (2010a, see Table 6) for which we have near-infrared observations of the cluster galaxies (Haines et al., 2009). As such, they were selected without reference to their X-ray morphology and temperature structure, and yielded a satisfactory weak-shear profile fit to a Navarro et al. (1997) density profile.

[1] http://www.sr.bham.ac.uk/locuss

Table 3.1: Cluster sample

Name	Redshift	RA [J2000]	Dec [J2000]	N_{gal} (<1Mpc)	Completeness (% <1Mpc)	$M_{2D}(<1\mathrm{Mpc})$ ($10^{14}M_\odot$)	$L_K(<1\mathrm{Mpc})$ ($10^{12}L_\odot$)	M_{500} ($10^{14}M_\odot$)	$L_K(<r_{500})$ ($10^{12}L_\odot$)
ABELL0068	0.2546	00 37 05.28	+09 09 10.8	49	58	$7.66^{+2.17}_{-2.17}$	$13.45^{+2.19}_{-2.19}$	$4.17^{+1.23}_{-1.07}$	$13.79^{+2.61}_{-2.53}$
ABELL0115a	0.1971	00 55 59.76	+26 22 40.8	65	73	$9.93^{+3.49}_{-3.49}$	$14.34^{+2.22}_{-2.22}$	$3.86^{+1.64}_{-1.33}$	$14.98^{+3.07}_{-2.81}$
ABELL0209	0.2060	01 31 53.00	−13 36 34.0	99	80	$13.04^{+1.46}_{-1.46}$	$19.73^{+2.12}_{-2.12}$	$8.84^{+1.36}_{-1.23}$	$26.49^{+2.78}_{-2.72}$
RXJ0142.0+2131	0.2803	01 42 02.64	+21 31 19.2	57	67	$7.87^{+1.93}_{-1.93}$	$14.24^{+1.98}_{-1.98}$	$4.07^{+0.86}_{-0.76}$	$15.05^{+2.30}_{-2.24}$
ABELL0267	0.2300	01 52 48.72	+01 01 08.4	25	31	$6.74^{+1.44}_{-1.44}$	$12.40^{+4.82}_{-4.82}$	$3.30^{+0.69}_{-0.61}$	$12.40^{+4.92}_{-4.90}$
ABELL0291	0.1960	02 01 44.20	−01 12 03.0	42	61	$7.55^{+1.56}_{-1.56}$	$10.18^{+1.44}_{-1.44}$	$4.11^{+1.00}_{-0.89}$	$10.23^{+1.68}_{-1.63}$
ABELL0383	0.1883	02 48 02.00	−03 32 15.0	56	87	$7.59^{+1.61}_{-1.61}$	$9.10^{+1.79}_{-1.79}$	$3.39^{+0.73}_{-0.61}$	$9.75^{+1.96}_{-1.92}$
ABELL0586	0.1710	07 32 22.32	+31 38 02.4	76	71	$10.78^{+3.46}_{-3.46}$	$20.92^{+4.27}_{-4.27}$	$6.77^{+2.00}_{-1.63}$	$25.83^{+5.37}_{-5.16}$
ABELL0611	0.2880	08 00 55.92	+36 03 39.6	64	72	$10.22^{+1.94}_{-1.94}$	$16.95^{+3.70}_{-3.70}$	$5.19^{+1.00}_{-0.91}$	$19.00^{+3.94}_{-3.91}$
ABELL0697	0.2820	08 42 57.84	+36 21 54.0	77	83	$11.91^{+1.62}_{-1.62}$	$16.17^{+3.44}_{-3.44}$	$8.39^{+1.27}_{-1.17}$	$22.09^{+4.01}_{-3.99}$
ABELL1835	0.2528	14 01 02.40	+02 52 55.2	127	91	$15.70^{+2.94}_{-2.94}$	$22.89^{+2.96}_{-2.96}$	$9.69^{+1.71}_{-1.53}$	$28.04^{+3.46}_{-3.37}$
ZwCl1454.8+2233	0.2578	14 57 14.40	+22 20 38.4	40	78	$7.07^{+2.89}_{-2.89}$	$8.68^{+2.33}_{-2.33}$	$2.61^{+0.99}_{-0.81}$	$8.25^{+2.54}_{-2.47}$
ABELL2219	0.2281	16 40 22.56	+46 42 21.6	113	80	$11.10^{+2.26}_{-2.26}$	$21.42^{+2.79}_{-2.79}$	$8.10^{+1.50}_{-1.36}$	$26.69^{+3.43}_{-3.36}$
RXJ1720.1+2638	0.1640	17 20 08.88	+26 38 06.0	70	98	$6.17^{+2.02}_{-2.02}$	$10.44^{+2.05}_{-2.05}$	$3.77^{+1.11}_{-0.94}$	$10.79^{+2.32}_{-2.24}$
RXJ2129.6+0005	0.2350	21 29 37.92	+00 05 38.4	40	70	$8.37^{+1.83}_{-1.83}$	$9.21^{+2.57}_{-2.57}$	$4.69^{+1.10}_{-0.99}$	$10.73^{+2.77}_{-2.74}$
ABELL2390	0.2329	21 53 36.72	+17 41 31.2	122	85	$13.75^{+1.99}_{-1.99}$	$20.05^{+2.02}_{-2.02}$	$7.10^{+1.29}_{-1.17}$	$21.66^{+2.45}_{-2.39}$
ABELL2485	0.2472	22 48 31.13	−16 06 25.6	51	85	$7.74^{+2.39}_{-2.39}$	$10.09^{+2.48}_{-2.48}$	$3.29^{+0.90}_{-0.80}$	$9.81^{+2.63}_{-2.60}$

N_{gal}: Number of spectroscopically confirmed member galaxies with $K \leq K^*(z) + 1.5$. Completeness: Percentage of galaxies with $K \leq K^*(z) + 1.5$ and within the $J-K$ colour cut that have spectroscopic data.

3.2.2 Gravitational Weak-Lensing Masses

We use both model independent projected and model dependent deprojected weak-lensing masses from Okabe et al. (2010a) (Table 3.1), in which Subaru/Suprime-Cam imaging was used to map the distribution of matter in each cluster. Details of the weak-lensing analysis can be found in Okabe et al. and are summarised here. Using deep V- and i'-band data, background galaxies were selected as those redder or bluer than the cluster red sequence (following Umetsu & Broadhurst, 2008; Umetsu et al., 2009), and their redshifts estimated statistically by matching their colours and magnitudes to the COSMOS photometric redshift catalogue (Ilbert et al., 2009). The KSB method (Kaiser et al., 1995) was used to measure a shear estimate for each galaxy, by considering the point spread function (PSF) and residual mean ellipticity of point sources.

The model independent mass is estimated using aperture mass densitometry, as the azimuthally averaged tangential shear is related to the projected mass density. The ζ_c-statistic (Clowe et al., 2000) relates the tangential shear to the 2D mass enclosed within a circular aperture. The 3D spherical mass, M_Δ, is defined as the mass within radius r_Δ, the radius within which the average density is $\Delta \times \rho_{\rm crit}$, where $\rho_{\rm crit} = 3H(z)^2/8\pi G$, the critical density of the Universe. The values for M_Δ are estimated by fitting to the measured shear profile an NFW model parameterised by M_Δ and c_Δ (the concentration parameter), where $\rho(r) \propto (c_\Delta r/r_\Delta)^{-1}(1 + c_\Delta r/r_\Delta)^{-2}$ (Navarro et al., 1997).

We also consider the 3D spherical mass within a fixed radius, and the projected mass within the r_Δ values determined by the 3D analysis. We work with an overdensity $\Delta = 500$ as r_{500} is typically the limiting radius to which all mass measurement methods can probe, enabling comparisons, and a fixed radius of 1Mpc because $r_{500} \simeq 1$Mpc for our sample.

Recent results (Planck Collaboration et al., 2013; Okabe et al., 2013; Applegate et al., 2014) suggest that Okabe et al.'s (2010a) M_{500} values may be underestimated by up to 20%, with no obvious trend with mass. We therefore concentrate on the slope and scatter of the mass-luminosity relation. We will consider the absolute normalisation of the mass-luminosity relation and explore possible subtle systematics in the scatter and slope of the relation in Chapter 5.

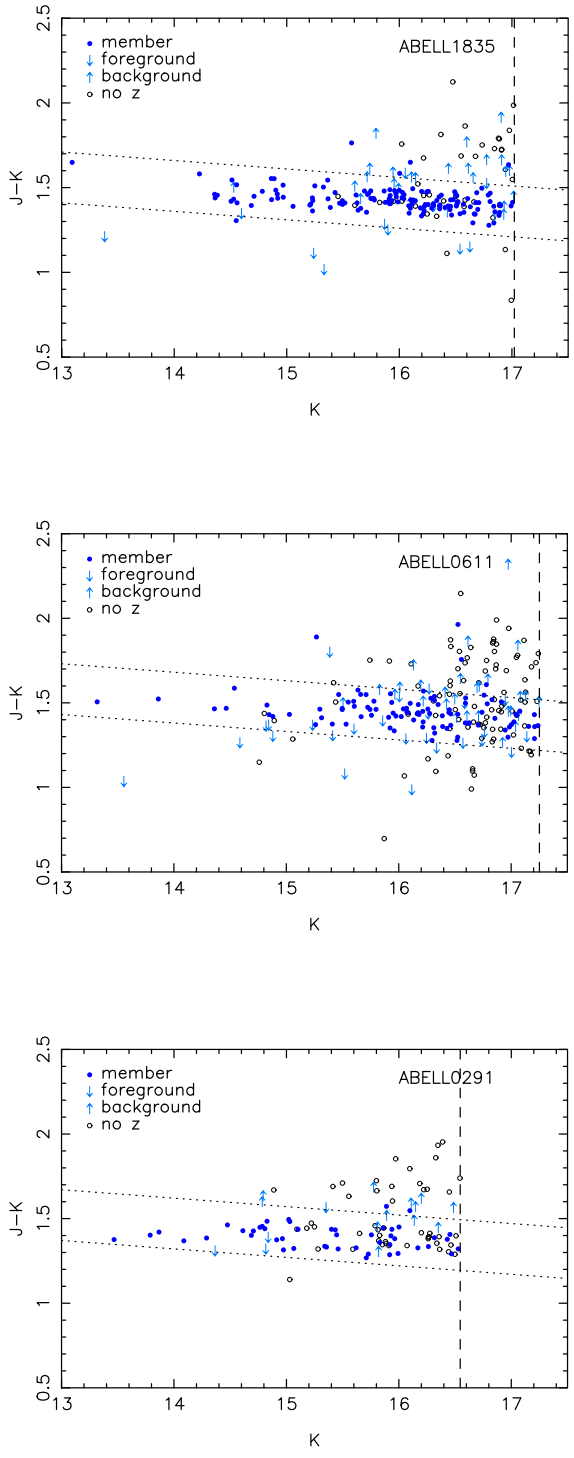

Figure 3.1: Colour-magnitude plots for three example clusters: the cluster with the highest number of galaxies above the magnitude cut $K^*(z) + 1.5$ - ABELL1835, the middle - ABELL0611, and the lowest - ABELL0291. The well defined ridge line of confirmed cluster members (dark blue filled points) can be clearly seen. Up and down arrows show background and foreground galaxies respectively, and hollow points show the galaxies with no spectroscopic data. The dotted lines show the width of the colour cut used for the colour selected L_K measurements, and the vertical dashed lines mark $K^*(z) + 1.5$ for the respective cluster redshifts.

3.2.3 Observations

We have observed 15 clusters from our sample with WFCAM on UKIRT, and the remaining two clusters with NEWFIRM on the Mayall 4-m telescope at Kitt Peak National Observatory. Details of these observations can be found in Haines et al. (2009) and are summarised here. The WFCAM data cover $52' \times 52'$ fields of view, while the NEWFIRM data consist of dithered and stacked images covering $27' \times 27'$ fields of view, both to depths of $K \simeq 19$, $J \simeq 21$ with FWHM $\simeq 1''$. Total K band Kron magnitudes were determined for each source, while $(J-K)$ colours were derived within fixed circular apertures of diameter $2''$. The typical error on an individual galaxy magnitude is a few per cent.

Galaxy colours can be difficult to interpret, particularly in the optical, because they are affected by redshift, metallicity, star formation rate and dust extinction. However, near-infrared wavelengths are relatively insensitive to the latter two ($A_V/A_K \sim 10$, Fitzpatrick, 1999), while $(J-K)$ evolves monotonically with redshift out to $z \sim 0.5$. This means there is no distinction between the red sequence and the blue cloud; galaxies of a particular redshift lie along a single narrow relation in the $(J-K)/K$ colour-magnitude diagram (Figure 3.1), allowing us to simply select galaxies within a colour slice around this sequence in order to select all galaxies (passive and star forming) within a redshift range centred on the cluster. This is in contrast to optical colour-magnitude diagrams which show a prominent blue cloud (e.g. Baldry et al., 2004), and a larger range of deviations from the red sequence within the cluster member population.

In addition to near-infrared data, we have spectroscopic data from MMT/Hectospec, observed as part of the Arizona Cluster Redshift Survey (ACReS[2]; M. J. Pereira et al. in preparation). The observation details can be found in Haines et al. (2013) and are summarised here. Hectospec is a 300-fibre multi-object spectrograph with a field of view of 1° diameter on the 6.5m MMT telescope. The 270 line grating was used, providing a wide wavelength range (3650–9200Å) at 6.2Å resolution. Redshifts were determined by comparison of the reduced spectra with stellar, galaxy and quasar template spectra. Galaxies that fall within a colour slice around the ridge line of cluster members in the $(J-K)/K$ colour-magnitude diagram (Figure 3.1) were targeted by ACReS.

[2]http://herschel.as.arizona.edu/acres/acres.html

3.2.4 Near-Infrared Luminosity

As with the mass measurements we calculate luminosities within both 1Mpc and r_{500}.

To determine which galaxies are in a cluster we plot their redshifts against distance from the centre of the cluster, which shows a trumpet shaped caustic profile as expected for galaxies infalling and orbiting within a massive gravitational structure. All galaxies within this caustic are identified as cluster members, and we select all those within a circular aperture (of radius r_{500}, and 1Mpc) on the sky. To account for spectroscopic incompleteness we weight each galaxy by the inverse probability of it having been observed spectroscopically. We give an initial equal weight (1.0) to all those galaxies which could have been targeted for spectroscopy. For each galaxy lacking a redshift, its weight is transferred equally to its ten nearest neighbouring galaxies on the sky with known redshift that had the same priority level in the targeting strategy.

Due to the magnitude limit of the spectroscopic coverage we only consider galaxies with $K \leq K^*(\mathrm{z}) + 1.5$, for which the average spectroscopic completeness is 75% within 1Mpc (Table 3.1). We base our estimates of $K^*(\mathrm{z})$ on Lin et al. (2006). To convert from apparent K band magnitude to rest frame luminosity, we use a k-correction consistent with Mannucci et al. (2001), and the absolute K band magnitude of the sun, $M_{K,\odot} = 3.39$ (Johnson, 1966). To account for the contribution of faint galaxies with $K > K^*(\mathrm{z}) + 1.5$ we multiply the cluster luminosities by a factor of 1.286, calculated by assuming that the faint end of the cluster galaxy luminosity function has a slope of $\alpha = -1.0$ (e.g. Balogh et al., 2011).

We also use a second method to calculate cluster luminosity, which differs only in how cluster membership is determined. Spectroscopic data will not necessarily be available for large samples in future surveys, and so instead we use the $(J-K)/K$ colour-magnitude plots. Probable cluster members are identified as those lying within ± 0.15mags of the ridge line of cluster members in the $(J-K)/K$ colour-magnitude plots (Figure 3.1), and the luminosity calculation continues as above. We carry out a statistical background correction using two control fields - the UKIDSS-DXS Lockman Hole and XMM-LSS fields (Lawrence et al., 2007). For each cluster we place 30 apertures of radius matching that used for the cluster luminosity measurements, and perform the same colour selection and luminosity calculation. The mean and standard deviation on the background calculated in this way are subtracted from our cluster luminosity measurements and propagated into the error respectively. The colour selection identifies all but 48 ($< 3\%$) of the

Table 3.2: Scaling relation parameters

Member Selection	Radius	Normalisation a	Slope b	Intrinsic Scatter $\sigma_{\ln M_{\rm WL}\|L_K}$, %
Model independent projected mass				
Spectroscopic	1Mpc	$1.06^{+0.98}_{-0.58}$	$0.83^{+0.27}_{-0.24}$	10^{+8}_{-5}
Spectroscopic	r_{500}	$0.52^{+0.40}_{-0.25}$	$0.96^{+0.22}_{-0.20}$	11^{+8}_{-6}
Colour	1Mpc	$0.84^{+1.02}_{-0.53}$	$0.96^{+0.38}_{-0.30}$	11^{+8}_{-6}
Colour	r_{500}	$0.42^{+0.41}_{-0.22}$	$1.10^{+0.27}_{-0.25}$	13^{+9}_{-7}
Model dependent deprojected mass				
Spectroscopic	1Mpc	$0.77^{+0.49}_{-0.35}$	$0.80^{+0.23}_{-0.18}$	8^{+6}_{-4}
Spectroscopic	r_{500}	$0.35^{+0.22}_{-0.14}$	$0.97^{+0.17}_{-0.17}$	10^{+7}_{-5}
Spectroscopic	r_{200}	$0.44^{+0.38}_{-0.19}$	$0.92^{+0.20}_{-0.19}$	16^{+9}_{-8}
Colour	1Mpc	$0.62^{+0.51}_{-0.33}$	$0.93^{+0.29}_{-0.24}$	10^{+7}_{-5}
Colour	r_{500}	$0.27^{+0.22}_{-0.14}$	$1.12^{+0.26}_{-0.22}$	13^{+8}_{-6}
Colour	r_{200}	$0.25^{+0.29}_{-0.13}$	$1.07^{+0.24}_{-0.24}$	16^{+10}_{-8}
Spectroscopic	L : 1Mpc, M : r_{500}	$0.19^{+0.21}_{-0.11}$	$1.24^{+0.32}_{-0.27}$	13^{+9}_{-7}
Colour	L : 1Mpc, M : r_{500}	$0.13^{+0.22}_{-0.09}$	$1.44^{+0.46}_{-0.45}$	15^{+11}_{-8}

confirmed members of the entire sample.

The error on the luminosity for each cluster is calculated from several components added in quadrature. The first, bootstrap resampling with replacement, involves calculating the cluster luminosity for 10^5 resamples of its members, and the standard deviation of these luminosities is the error contribution. Another component, which is only valid for $L(< r_\Delta)$, comes from the uncertainty in the radius, which comes from the uncertainty in the mass and causes an error in the luminosity.

The average of the ratio of luminosities calculated using both methods, $\langle L_{\rm spect.}/L_{\rm colour}\rangle$, is 0.97 ± 0.06 within r_{500} and 0.98 ± 0.06 within 1Mpc; the consistency with unity showing the consistency between the methods on average.

3.3 Results

In this section we model the relation between mass and K band luminosity. We measure the quantities within both 1Mpc and r_{500}, consider both 2D projected and 3D deprojected masses, and use luminosities based on both spectroscopic and colour member selection.

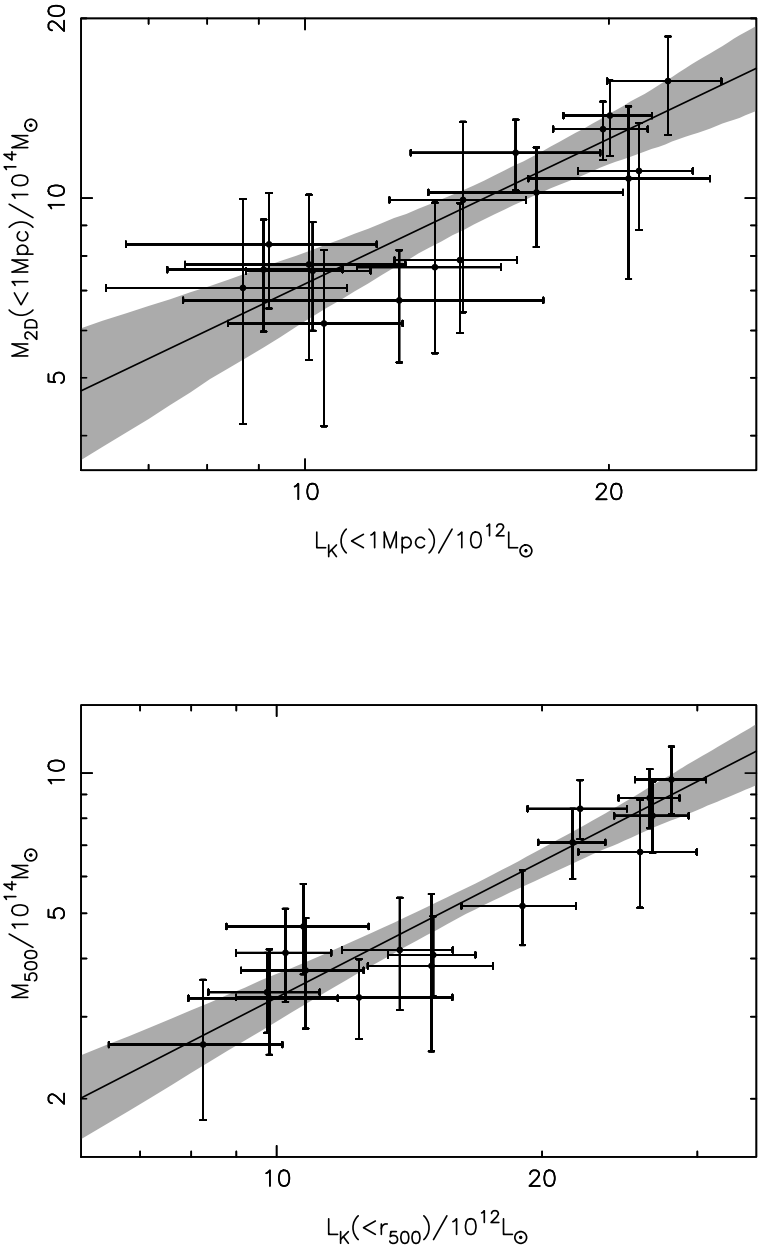

Figure 3.2: Scaling relations between weak-lensing mass and spectroscopically confirmed near-infrared luminosity, where we show the data points, resultant scaling relation and 68 per cent confidence region. Upper: the relation between projected mass and luminosity within a fixed metric aperture of 1Mpc. Lower: the relation between the deprojected 3D mass and luminosity within r_{500}.

3.3.1 Fitting Method

To analyse the scaling relation between $M_{\rm WL}$ and L_K we linearise the problem by taking the base-10 log of the respective measurements, and use a Bayesian approach to linear regression with a publicly available IDL code (Kelly, 2007). Kelly highlights the importance of correctly handling measurement errors when performing linear regression, and demonstrates that this model outperforms other estimators (OLS, BCES, FITEXY), especially when the measurement errors are large. The result of the routine is a line of best fit of the form:

$$\frac{M_{\rm WL}}{10^{14}M_{\odot}} = a\left(\frac{L_K}{10^{12}L_{\odot}}\right)^b, \tag{3.1}$$

with normalisation a, slope b, and intrinsic scatter $\sigma_{\ln M_{\rm WL}|L_K}$.

3.3.2 $M_{\rm WL} - L_K$ Relation

We first consider the relation between the 2D projected mass and luminosity within 1Mpc, because these quantities can be calculated directly from the data, with the fewest assumptions. Importantly, the use of a fixed metric aperture guarantees that the covariance between the mass and luminosity is zero. We find a slope of $b = 0.83^{+0.27}_{-0.24}$ and an intrinsic scatter of $\sigma = 10^{+8}_{-5}\%$ (Figure 3.2, Table 3.2).

The most common mass studied in the literature is the 3D overdensity mass M_{Δ}. We therefore also consider the scaling relation between deprojected mass and luminosity within r_{500}, both to enable comparisons with the literature, and because the halo mass function is typically expressed in terms of M_{Δ}. We find the relation between 3D deprojected mass and luminosity within r_{500} is parameterised by $b = 0.97^{+0.17}_{-0.17}$ and $\sigma = 10^{+7}_{-5}\%$ (Figure 3.2, Table 3.2), again showing a promising low scatter.

We note that measuring the deprojected mass and luminosity within radii that scale with mass inevitably introduces covariance between the variables. In general, covariance may suppress the measured scatter in scaling relations. However we draw attention to the consistency between the scatter measured for the relation between quantities inside a fixed metric aperture, and the result within r_{500} above. This indicates that the impact of the covariance on the measured scatter is negligible.

From a cosmological perspective, the most meaningful mass measurement is the deprojected

spherical mass M_{500}, however from an observational perspective, the simplest luminosity to measure is $L(< 1\text{Mpc})$. We therefore fit a relation between these two values, finding $b = 1.24^{+0.32}_{-0.27}$ and $\sigma = 13^{+9}_{-7}$ (Table 3.2). This relation is particularly important in demonstrating the potential of L_K as a mass proxy for cluster cosmology, as measuring $L(< 1\text{Mpc})$ does not require any prior radial information.

For completeness, we also measure the relations between deprojected mass and luminosity within 1Mpc, and projected mass and luminosity within r_{500}. We find that these relations also have low scatter, of $\sigma = 8^{+6}_{-4}\%$ and $\sigma = 11^{+8}_{-6}\%$ respectively, and that the slope of relations based on projected and deprojected mass are in close agreement (Table 3.2). Indeed, the slope of all of the spectroscopic relations is consistent with unity, and in agreement within the errors.

However, we note that the central value of the slope of relations calculated within 1Mpc are consistently shallower than those calculated within r_{500}. Previous observational studies (e.g. Carlberg et al., 1997b; Lin et al., 2004; Hansen et al., 2005) have shown that the number density profile of cluster galaxies is fit well by an NFW distribution, and Budzynski et al. (2012) showed that the concentration parameter for the number density profile is a factor of two smaller than that of the dark matter density profile. This causes the stellar fraction to increase with cluster radius, following the same trend for all clusters relative to the overdensity radius. The fixed radius corresponds to a higher overdensity radius in larger clusters, and so results in a decreased stellar fraction, while the opposite is true for smaller clusters. This steepens the $L_K/M - M$ relation at 1Mpc compared to r_{500}, which leads to a shallower $M - L_K$ relation at 1Mpc compared to r_{500}.

Finally, we fit the scaling relation model to the same weak-lensing masses as discussed above, and near-infrared luminosities that are based on colour selection, as described in Section 3.2.4. We find that these colour selected scaling relations are fully consistent with the spectroscopically confirmed relations (Table 3.2).

Table 3.3: Comparison with literature

Paper	Sample Size	Mass Measurement Technique	Mass Range $10^{14} M_\odot$	Redshift Range	Slope b	Intrinsic Scatter $\sigma_{\ln M\|L_K}$, %
			M_{500}			
Balogh et al. (2011)	13	X-Ray	$0.55 \leq M_{500} \leq 7.06$	$0.05 \leq z \leq 0.095$	$1.30^{+0.45}_{-0.46}$	64^{+22}_{-16}
Lin et al. (2004)	93	$M - T_X$ Relation	$0.2 \leq M_{500} \leq 12.6$	$0.016 \leq z \leq 0.09$	$1.18^{+0.07}_{-0.07}$	25^{+6}_{-5}
Lin et al. (2003)	27	$M - T_X$ Relation	$0.78 \leq M_{500} \leq 8.3$	$0.016 \leq z \leq 0.09$	$1.00^{+0.16}_{-0.16}$	28^{+9}_{-7}
This work	17	Weak-Lensing	$2.6 \leq M_{500} \leq 9.7$	$0.16 \leq z \leq 0.29$	$0.99^{+0.21}_{-0.18}$	11^{+8}_{-6}
			M_{200}			
Ramella et al. (2004)[b]	55	Velocity Dispersion	$0.007 \leq M_{200} \leq 10.23$	$z \leq 0.04$	$1.21^{+0.14}_{-0.14}$	57^{+10}_{-9}
Ramella et al. (2004)[c]	61	Velocity Dispersion	$0.007 \leq M_{200} \leq 15.49$	$z \leq 0.05$	$1.21^{+0.09}_{-0.09}$	50^{+9}_{-8}
Rines et al. (2004)	9	Caustics	$0.76 \leq M_{200} \leq 7.8$	$z \leq 0.05$	$1.17^{+0.36}_{-0.30}$	29^{+21}_{-15}
Ramella et al. (2004)[a]	36	Velocity Dispersion	$0.039 \leq M_{200} \leq 10.23$	$z \leq 0.04$	$1.12^{+0.25}_{-0.25}$	62^{+13}_{-12}
Muzzin et al. (2007)	14	Velocity Dispersion	$3.5 \leq M_{200} \leq 33.3$	$0.17 \leq z \leq 0.54$	$1.08^{+0.29}_{-0.29}$	45^{+17}_{-12}
Lin et al. (2004)	93	$M - T_X$ Relation	$0.3 \leq M_{200} \leq 18.9$	$0.016 \leq z \leq 0.09$	$1.07^{+0.06}_{-0.06}$	28^{+5}_{-4}
This work	17	Weak-Lensing	$4.0 \leq M_{200} \leq 15.5$	$0.16 \leq z \leq 0.29$	$0.93^{+0.20}_{-0.19}$	16^{+10}_{-8}
Balogh et al. (2011)[d]	18	Velocity Dispersion	$1.66 \leq M_{200} \leq 5.97$	$0.05 \leq z \leq 0.096$	$0.05^{+0.18}_{-0.17}$	24^{+12}_{-10}

Properties of samples of groups/clusters with M and L_K values in the literature. [a]Core sample. [b]Total sample. [c]Extended sample, including 5 Rines et al. (2004) groups/clusters. [d]The shallow slope of this sample is likely a consequence of limiting the dynamic range in the dynamical mass, as noted by the authors.

3.4 Discussion

In Section 3.4.1 we compare our results with other weak-lensing based mass-observable scaling relations, and in Section 3.4.2 we compare our results with previous measurements of the scaling relation between mass and near-infrared luminosity.

3.4.1 Comparison with previous weak-lensing based scaling relation results

Our results, based on a small pilot study sample, show that the intrinsic scatter in the scaling relation between weak-lensing mass and near-infrared luminosity is $\simeq 10\%$ on scales of 1Mpc, which corresponds to an overdensity of 500 with respect to the critical density of the Universe. This result is independent of whether the scaling relation is derived from measurements within a fixed metric aperture, or within a radius (r_{500}) that scales with mass, and independent of whether the luminosity is based on spectroscopically confirmed members or galaxies selected in the $(J-K)/K$ colour-magnitude plane. The scatter in weak-lensing mass to near-infrared luminosity scaling relation is therefore smaller than that found in all previous weak-lensing-based studies of mass-observable scaling relations (Okabe et al., 2010b; Marrone et al., 2012; Mahdavi et al., 2013), with the exception of Hoekstra et al.'s (2012) relation between mass and the integrated Compton Y_{SZ} parameter.

These results all point to observables that are closely related to a line-of-sight integral of a linear quantity through the cluster potential being low scatter proxies for the weak-lensing mass of clusters. Arguably the projected near-infrared luminosity of a cluster within a fixed metric aperture is the least expensive and least model dependent of the available observables because it is based on simply measuring flux from galaxies above a well-defined limit, and is feasible with wide field survey data.

3.4.2 Comparison with previous studies of $M - L_K$

The intrinsic scatter in our $M_{500} - L_K$ relation is much lower than the scatter of $\sigma_{\ln L_K|M} = 28\%$ found by Lin et al. (2003). These authors estimated M_{500} from the relationship between hydrostatic mass and X-ray temperature. To compare our work more directly with Lin et al. we repeat our fit of the $M_{500} - L_K$ relation using hydrostatic masses (Martino et al., 2014) in place of our weak-lensing masses. We measure an intrinsic scatter of $\sigma_{\ln M|L_K} = 25^{+11}_{-10}\%$, which

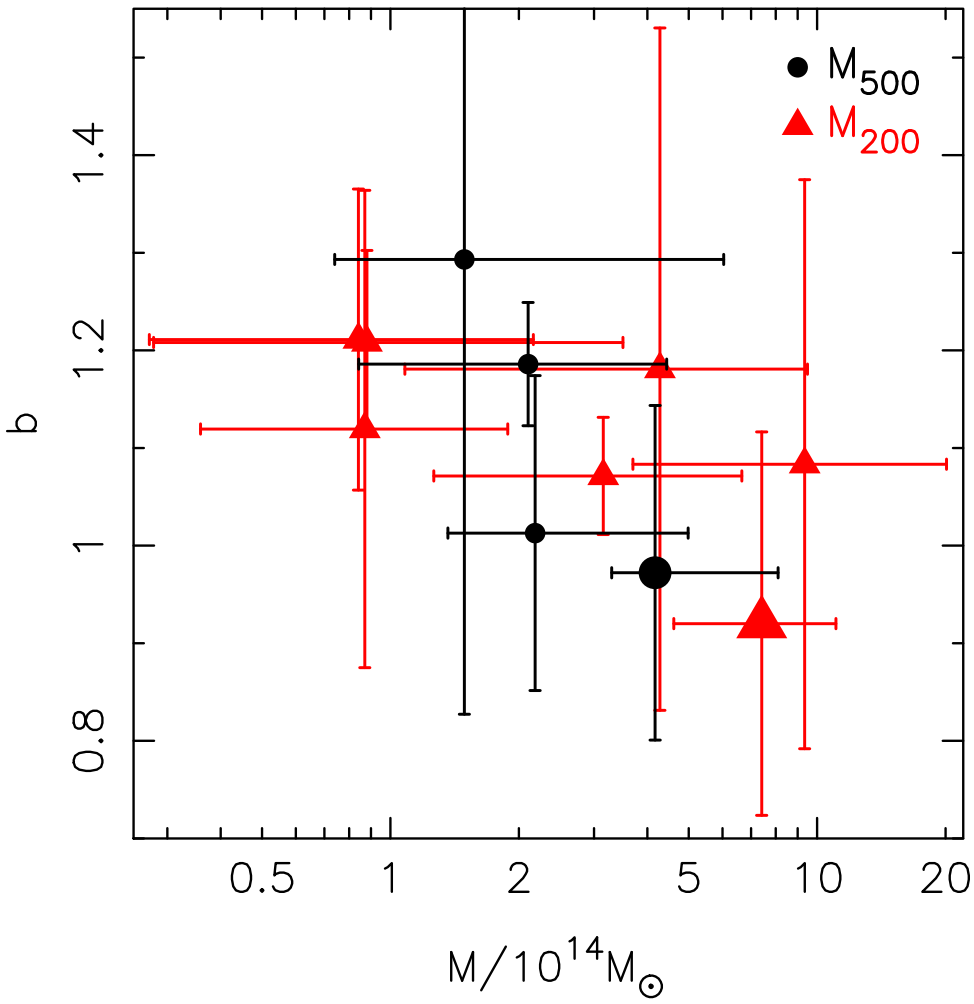

Figure 3.3: The slopes of the $M = aL_K^b$ relation fit to each literature sample using the Kelly (2007) method (Table 3.3). The points show the slope against the average mass, the vertical error bars show the error on the slope and the horizontal error bars enclose 68% of the mass range. The large points are the results from this work using spectroscopic member selection and 3D NFW masses. Note that the Balogh et al. (2011) $\Delta = 500$ data is not visible as the slope is much shallower than the other results.

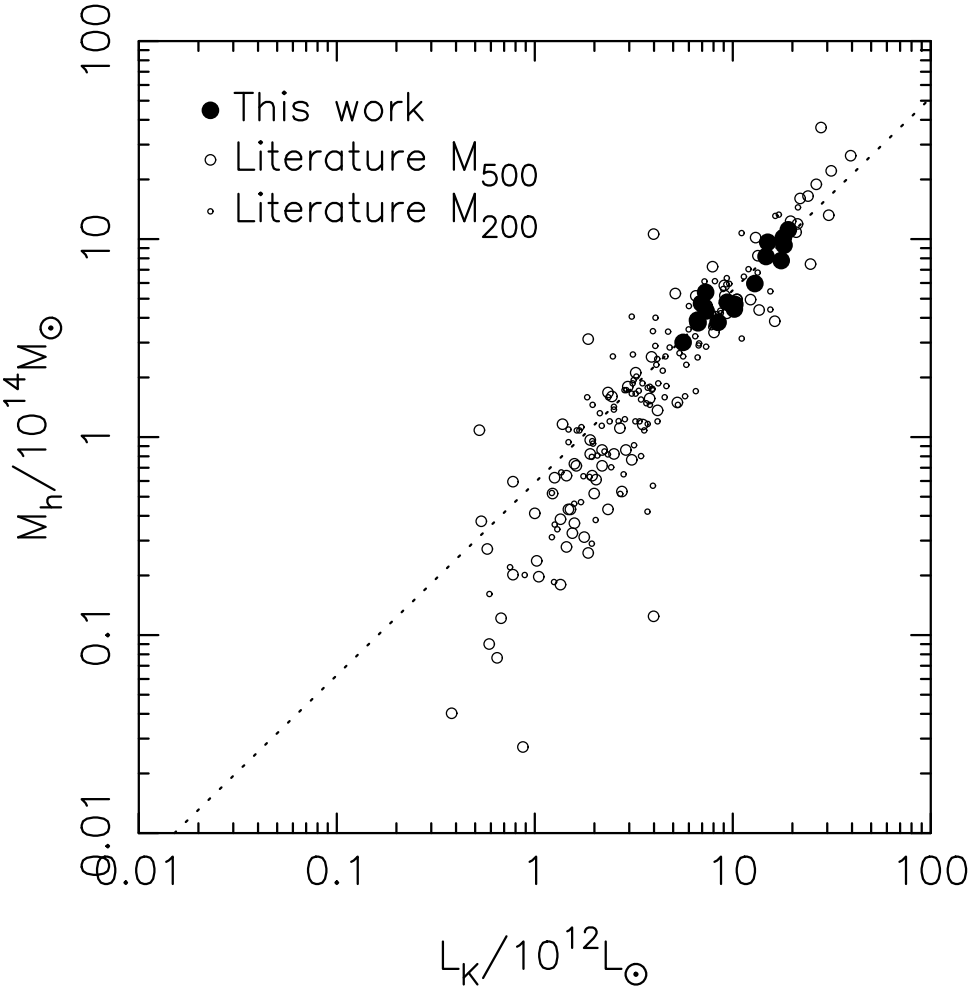

Figure 3.4: Comparison of all the available data from the literature, as summarised in Table 3.3, normalised such that the individual scaling relations overlap with our $M_{\mathrm{WL},500}$ relation (dotted line) at our mean M_{500} value. For the clusters analysed in two papers and/or at two overdensities, we plot the most recent and/or highest overdensity values, and use M_h to refer to the total halo mass.

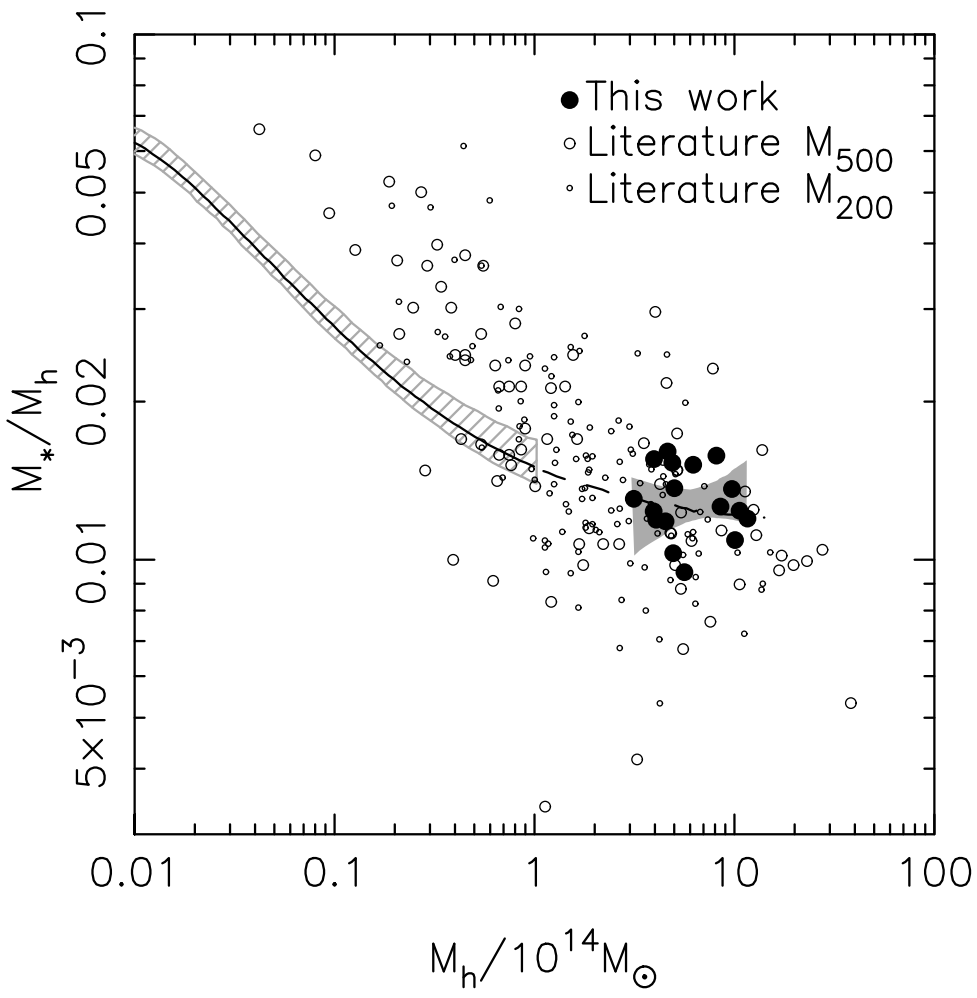

Figure 3.5: The data from Figure 3.4, where L_K has been converted to M_* using a mass to light ratio of 0.73 (Cole et al., 2001). Also shown is the trend and error envelope from Leauthaud et al.'s (2012) halo occupation distribution model within r_{500}; where the dashed line shows the extrapolation beyond the data. The solid shaded region shows the error envelope from our $M_{500} - L_K$ relation.

is consistent with Lin et al.'s result, and supports the interpretation of the weak-lensing based scaling relation results discussed in Section 3.4.1.

We now concentrate on comparing our $M_{WL} - L_K$ results with those in the literature, and make two corrections to ensure that our data are comparable. Firstly, we multiply our M_{500} values by 1.20, to account for the 20% bias in the mass measurements as discussed in Section 3.2.2. Secondly, we deproject our luminosities. Counting all the cluster members within r_Δ on the sky gives a cylindrical volume projected along the line of sight within which we calculate the luminosity, which requires deprojection to correct to a spherical volume. We therefore multiply our L_{500} values by 0.68, the average ratio of the 3D to 2D M_{500} measurements (Okabe et al., 2010a). We use a constant based on the NFW profile for this deprojection, as do Muzzin et al. (2007) (0.791), Ramella et al. (2004) (0.80) and Giodini et al. (2009) (0.86). After these corrections our error weighted mean mass-to-light ratio is $55.9 \pm 1.8 M_\odot/L_\odot$, which is consistent with other results in the literature (e.g. Rines et al., 2001; Kochanek et al., 2003; Lin et al., 2003; Rines et al., 2004; Muzzin et al., 2007).

Results in the literature are generally expressed as $L_K = a\,M^b$, and in that form the slope of our deprojected mass relation within r_{500} is $b = 1.00^{+0.21}_{-0.18}$. The published results are generally shallower than this (equivalent to steeper in the form $M = aL_K^b$). To ensure that this is not caused by a difference in fitting method, we refit each sample in the published literature with the Kelly (2007) method in the same manner as our results in Section 3.3, in the form $M = aL_K^b$ (Table 3.3). We find that the flatter slope of our $M - L_K$ relation is not an artefact of fitting method. However we note that in general the dynamic range of mass explored by other authors is wider than our own, and extends to lower masses. This suggests that the slope of the $M - L_K$ relation may be a function of halo mass (Figure 3.3).

To further illustrate this point, we plot all the available data from the literature after re-normalising it with respect to our own, as we are focussing on the slope of the relation. For each sample we calculate the normalisation required to make the relevant best fit scaling relation intersect our relation at the mean mass of our sample, and apply that normalisation adjustment to every cluster in that sample (Figure 3.4).

We caution that the general shallowing of the $M - L_K$ relation may be an artefact of selection and/or measurement biases at low mass. Nevertheless, taking the gradual shallowing at face

value corresponds to a smaller stellar fraction for larger clusters. To explore this further we use a simple method to calculate $f_* \equiv M_*/M_h$ for our sample and compilation from the literature using a stellar mass-to-light ratio of 0.73 (Cole et al., 2001). For comparison we show the results from Leauthaud et al.'s (2012) halo occupation distribution model within r_{500}, noting that the same trend is found using abundance matching techniques (e.g. Guo et al., 2010; Moster et al., 2013; Behroozi et al., 2013; Kravtsov et al., 2014). The decreasing stellar fraction seen in Figure 3.5 suggests a quenching of star formation in larger systems, which is consistent with results of other observational studies (e.g. Gonzalez et al., 2007; Laganá et al., 2011). We also note that our results on the slope of the mass-luminosity relation of clusters – i.e. a linear relation between weak-lensing mass and K band luminosity – suggest that for the most massive halos the relationship between stellar mass fraction and halo mass may be flatter than implied by an extrapolation of by Leauthaud et al.'s relation. This emphasises the importance of direct calibration of this relation, as highlighted recently by Kravtsov et al. (2014).

3.5 Summary

In this pilot study we have shown that K band luminosity is a promising low scatter proxy for weak-lensing mass, with an intrinsic scatter of $\sim 10\%$.

A useful mass proxy must be easy to measure, and so we have considered the values closest to the data plane - projected values within 1Mpc - and found a scatter of only $\sigma_{\ln M_{\rm WL}|L_K} = 10^{+8}_{-5}\%$, demonstrating the practical potential of the relation. We have also shown that having spectroscopic information is not required, as the scatter does not increase when determining cluster membership using the $(J-K)/K$ colour-magnitude diagram. It will not be practical to have such spectroscopic coverage for future surveys, and so this is an important result.

The halo mass function is typically expressed in terms of M_Δ, so it is also of interest to study the scaling relation between mass and luminosity estimated within the three-dimensional overdensity radius. We therefore considered the relation between deprojected M_{500} and $L_K(< r_{500})$ and found an intrinsic scatter of only $\sigma_{\ln M_{\rm WL}|L_K} = 10^{+7}_{-6}\%$. We also note that the invariance of the scatter between the relation measured within a fixed metric aperture and that measured within r_{500} indicates that the impact of covariance between mass and luminosity via the use of r_{500} in the latter relation has negligible effect on the measured scatter.

The above relation was motivated by the most useful mass quantity for cosmology, while the first relation we considered was motivated by the most practical luminosity to measure. We combined the advantages of both these relations by considering the relation between M_{500} and $L(< 1\mathrm{Mpc})$. The resulting low scatter of only $\sigma_{\ln M_{\mathrm{WL}}|L_K} = 13^{+9}_{-7}\%$ demonstrates the potential of L_K as a mass proxy for cluster cosmology. This highlights the importance of calibrating the relation as a function of both redshift and mass.

The studies in the literature against which we compared our results used a range of mass measurements and find consistently higher scatter than our $\sim 10\%$, suggesting that the low intrinsic scatter in the $M_{\mathrm{WL}}-L_K$ relation is related to both quantities suffering similar projection effects. When compared to the literature there appears to be a mass dependence in the slope of the relation; the slope of the $M - L_K$ relation appears to be a decreasing function of mass. This is equivalent to a stellar fraction M_*/M_h that decreases with increasing mass, suggesting a quenching of star formation in larger systems.

Encouraged by the positive result of this pilot study, in Chapter 5 we will investigate this relation for a statistically complete sample of 42 clusters for which we now have near-infrared data. With this larger sample, and improved weak-lensing masses, we will be able to reduce statistical errors and subtle biases in our results and also investigate the effect of cluster morphology on the relation. We expect that our results will be helpful for upcoming large-scale optical/infrared surveys that will study galaxy clusters, with cosmological goals, including HSC, DES, Euclid, and LSST.

Chapter 4

Galaxy Cluster Luminosities and Colours, and their Dependence on Mass and Merger State

This chapter is taken from Mulroy et al. submitted: "Galaxy Cluster Luminosities and Colours, and their Dependence on Mass and Merger State", and includes $c_{\rm SB}$ measurements and values calculated from the Millennium simulation by co-authors.

4.1 Introduction

The composition of galaxy clusters is thought to represent that of the whole Universe, and so they offer a window into astrophysics on both cluster and galaxy scales (e.g. Kravtsov & Borgani, 2012). Their position at the extreme end of the mass function makes them sensitive to the underlying cosmology and provides a late time estimate of the cosmological parameters, complementary to alternative probes such as the cosmic microwave background and supernovae (e.g. Weinberg et al., 2013).

Accurate mass measurements of galaxy clusters are necessary to constrain the mass function, and thus cosmology (e.g. Allen et al., 2011). Methods to make such measurements include: dynamical, which measure the depth of the potential well of the clusters using the velocities of the galaxies; hydrostatic, which assume that the gas pressure is balanced by the gravitational attraction; and gravitational weak-lensing, which measure the distortion of the light distribution from distant galaxies by the gravitational potential of the cluster.

While these methods each have different biases that require further exploration, well-constrained

direct individual mass measurements require deep observations and extensive analysis that is not easily extended to very large samples. This motivates research into well calibrated scaling relations between easily measured 'mass proxies' and cluster mass. The preferable scaling relation is one with minimal intrinsic scatter between observable and mass, and an observable that is easily obtainable from survey data.

Potential observables that could be suitable mass proxies cover a wide range of the electromagnetic spectrum, including: millimetre Sunyaev Zel'dovich effect (e.g. Arnaud et al., 2010; Marrone et al., 2012), near-infrared luminosities (e.g. Lin et al., 2003; Mulroy et al., 2014), optical measures such as richness (e.g. Rozo et al., 2009; Andreon & Hurn, 2010) and velocity dispersion (e.g. Carlberg et al., 1997a; Ruel et al., 2014), and X-ray observables (e.g. Vikhlinin et al., 2006; Mantz et al., 2016b).

Promisingly, in Chapter 3 (Mulroy et al., 2014) we showed the total cluster near-infrared luminosity to be a low scatter mass proxy for a sample of clusters at $z \sim 0.23$. Future wide field surveys will observe clusters at higher redshifts, where their rest frame optical light has been redshifted into the near-infrared filters. It is therefore important to determine whether the small scatter found in the near-infrared luminosity persists at bluer rest frame wavelengths.

Extending the study of total cluster luminosity to bluer bands also allows us to investigate the colour of a galaxy cluster. This colour corresponds to the average member galaxy colour, which in turn is an indicator of the age and metallicity of the stellar population within it. It has been known for some time that galaxies within galaxy clusters have old stellar populations, low current star formation rates (SFR), and are relatively metal rich (Nelan et al., 2005; von der Linden et al., 2010; Smith et al., 2012). This highlights the influence of environment on galaxy properties, and motivates investigation into the galaxy populations within clusters of different evolutionary stages and morphological states.

The state of the cluster can be probed through central cluster properties. For instance, one common indicator of disturbance in the X-ray is cool core strength – a measure of the rate of gas cooling in the centre of a cluster. A strong cool core suggests a more relaxed history (e.g. Poole et al., 2008; Rossetti & Molendi, 2010). Cluster mergers can disturb not only the cool cores but also the gravitational potential of a cluster, which can be seen in the dynamics of the cluster galaxies and probed through the bulk cluster properties (Dressler & Shectman, 1988;

Burns, 1998).

Here we combine weak-lensing mass measurements, which have been shown in simulations to be unbiased on average (Oguri & Hamana, 2011; Becker & Kravtsov, 2011; Bahé et al., 2012), with optical luminosities, which require only shallow imaging data. Optical luminosities have previously been shown to be good proxies for X-ray and dynamical mass measurements (e.g. Girardi et al., 2000; Popesso et al., 2005). We utilise highly complete spectroscopic redshift catalogues in order to isolate issues arising from selecting members in colour-magnitude space, caused by the sensitivity of galaxy colour to astrophysics (e.g. Lu et al., 2009; Castignani & Benoist, 2016). We use the same member selection for every waveband to provide a clean probe of the underlying cluster physics.

In this chapter we use a sample of 19 massive galaxy clusters to quantify the scaling relations between optical luminosities and weak-lensing mass, before investigating the trends between cluster colour and various indicators of the level of disturbance in these clusters. We introduce our data in Section 4.2, present our results in Section 4.3 and our interpretation of cluster colour trends in Section 4.4, and summarise in Section 4.5. All photometric measurements are in the AB system, and we assume $\Omega_{\mathrm{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In this cosmology, at the average cluster redshift, $\langle z \rangle = 0.23$, 1 arcsec corresponds to a projected physical scale of 3.67 kpc.

4.2 Data

4.2.1 Sample

The sample comprises 19 X-ray luminous galaxy clusters at $0.15 < z < 0.30$ (Table 4.1), which populate the overlap between three surveys: the Sloan Digital Sky Survey (SDSS[1]), a wide field photometric and spectroscopic survey; the Local Cluster Substructure Survey (LoCuSS[2]) "High-L_X" sample, 50 well studied clusters from the multiwavelength survey of X-ray luminous clusters at $0.15 < z < 0.30$; and the Arizona Cluster Redshift Survey (ACReS[3]), a spectroscopic survey of 30 clusters drawn from the full LoCuSS sample. The LoCuSS "High-L_X" sample was selected on X-ray luminosity and the ACReS clusters are a representative sub-sample, while the

[1] http://www.sdss.org/
[2] http://www.sr.bham.ac.uk/locuss
[3] http://herschel.as.arizona.edu/acres/acres.html

overlap with SDSS is determined only by sky coverage. Thus, the main physical selection is on the X-ray luminosity.

4.2.2 Cluster Luminosities

We have total J and K band Kron magnitudes for the cluster galaxies from LoCuSS, most from WFCAM on UKIRT, and two (ZwCl0857.9+2107 and Abell0963) from NEWFIRM on the Mayall 4-m telescope at Kitt Peak National Observatory (Haines et al., 2009; Mulroy et al., 2014).

All clusters in our sample also have SDSS Data Release 12 *ugriz* band photometry (Gunn et al., 1998; Doi et al., 2010; Alam et al., 2015), from which we use the 'modelmag' aperture magnitudes and 'cmodelmag' total magnitudes. The u band data with a magnitude limit of 22.0 is not deep enough to robustly measure the predominantly red cluster galaxies at these redshifts, so we discard this bluest band. All magnitudes are corrected for galactic extinction assuming the dust maps of Schlegel, Finkbeiner & Davis (1998).

To determine cluster membership we use spectroscopic information from MMT/Hectospec observations taken by ACReS (Haines et al., 2013, M. J. Pereira et al. in prep.). Cluster members are those galaxies within the characteristic cluster caustic in redshift-clustercentric radius space. The spectroscopic targeting was K band limited (independent of colour) with a resulting average completeness of $\sim 75\%$ for galaxies with $K < K^*(\mathrm{z}) + 1.5$ within 1Mpc, and a weighting system was used to account for those objects not observed. This is calculated by weighting every potential spectroscopic target galaxy equally, then redistributing the weight from each galaxy lacking a redshift equally to its ten nearest neighbours on the sky that had the same priority level in the original targeting strategy.

In Figure 4.1, we show SDSS colour-magnitude diagrams for a typical cluster in our sample (Abell0068). The spectroscopically confirmed cluster members are shown in red, and demonstrate the tight red sequence typical of massive galaxy clusters.

To convert the magnitudes to rest frame luminosities we apply k-corrections derived from the polynomial fitting functions of Chilingarian et al. (2010) and Chilingarian & Zolotukhin (2012), and normalise to solar luminosity (Blanton & Roweis, 2007). At our redshifts, these fitting functions have been shown to agree on average with spectral energy distribution (SED)

fitting programs (e.g. `K-correct`, Blanton & Roweis, 2007) to within ~ 0.02 mags across the full range of optical and near-infrared data we use. We calculate the total cluster luminosity in each of the six bandpasses ($L_{\rm g}, L_{\rm r}, L_{\rm i}, L_{\rm z}, L_{\rm J}, L_{\rm K}$) by summing the weighted luminosity of all cluster galaxies within a clustercentric radius derived from weak-lensing analysis (see Section 4.2.3) and with $K < K^*(\rm z) + 1.5$, resulting in a roughly stellar mass limited selection of member galaxies. Uncertainties on the luminosities consist of two terms - one calculated by propagating the uncertainty on the weak-lensing radii, and the other from bootstrap resampling of the member galaxy luminosities.

4.2.3 Cluster Masses

We use weak-lensing masses from Okabe & Smith (2016), where the authors used Subaru/Suprime-Cam imaging and fit an NFW (Navarro, Frenk & White, 1997) mass density profile to the weak shear profile of each cluster. M_Δ is the mass calculated within r_Δ, the radius within which the average density is $\Delta \times \rho_{\rm crit}$, where $\rho_{\rm crit} = 3H(z)^2/8\pi G$, the critical density of the Universe. We consider the overdensities $\Delta = \Delta_{\rm vir}, 500, 2500$. $\Delta_{\rm vir}$ is defined as $\Delta_{\rm vir} = 18\pi^2 + 82x - 39x^2$ where $x = \Omega_{\rm m}(z) - 1$ (Bryan & Norman, 1998), and is equal to ~ 120 at the average redshift of our sample. The weak-lensing error analysis is described fully in Section 3.1 of Okabe & Smith (2016), and includes shape noise, photometric redshift uncertainties, and uncorrelated large-scale structure.

4.3 Results

4.3.1 Scaling Relations

We quantify the scaling relations between cluster luminosities L and weak-lensing masses $M_{\rm WL}$ by performing linear regression on the logarithmic values using the method of Kelly (2007). The scaling relation is parameterised as:

$$\frac{L}{10^{12} L_\odot} = a \left(\frac{M_{\rm WL}}{10^{15} M_\odot} \right)^b , \tag{4.1}$$

with intercept a, slope b, and intrinsic scatter $\sigma_{\ln L|M_{\rm WL}}$.

Table 4.1: Cluster sample

Cluster	RA α [J2000]	Dec δ [J2000]	Redshift z
Abell0068	9.2785	9.1566	0.2546
ZwCl0104.4+0048	16.7057	1.0564	0.2545
Abell0267	28.1748	1.0072	0.2300
Abell0291	30.4296	-2.1966	0.1960
Abell0586	113.0845	31.6335	0.1710
Abell0611	120.2367	36.0566	0.2880
Abell0697	130.7398	36.3666	0.2820
ZwCl0857.9+2107	135.1536	20.8946	0.2347
Abell0963	154.2652	39.0470	0.2060
Abell1689	197.8730	-1.3410	0.1832
Abell1758N	203.1600	50.5600	0.2792
Abell1763	203.8337	41.0012	0.2279
Abell1835	210.2588	2.8786	0.2528
Abell1914	216.4860	37.8165	0.1712
ZwCl1454.8+2233	224.3131	22.3428	0.2578
Abell2219	250.0827	46.7114	0.2281
RXJ1720.1+2638	260.0420	26.6260	0.1640
RXJ2129.6+0005	322.4165	0.0894	0.2350
Abell2390	328.4034	17.6955	0.2329

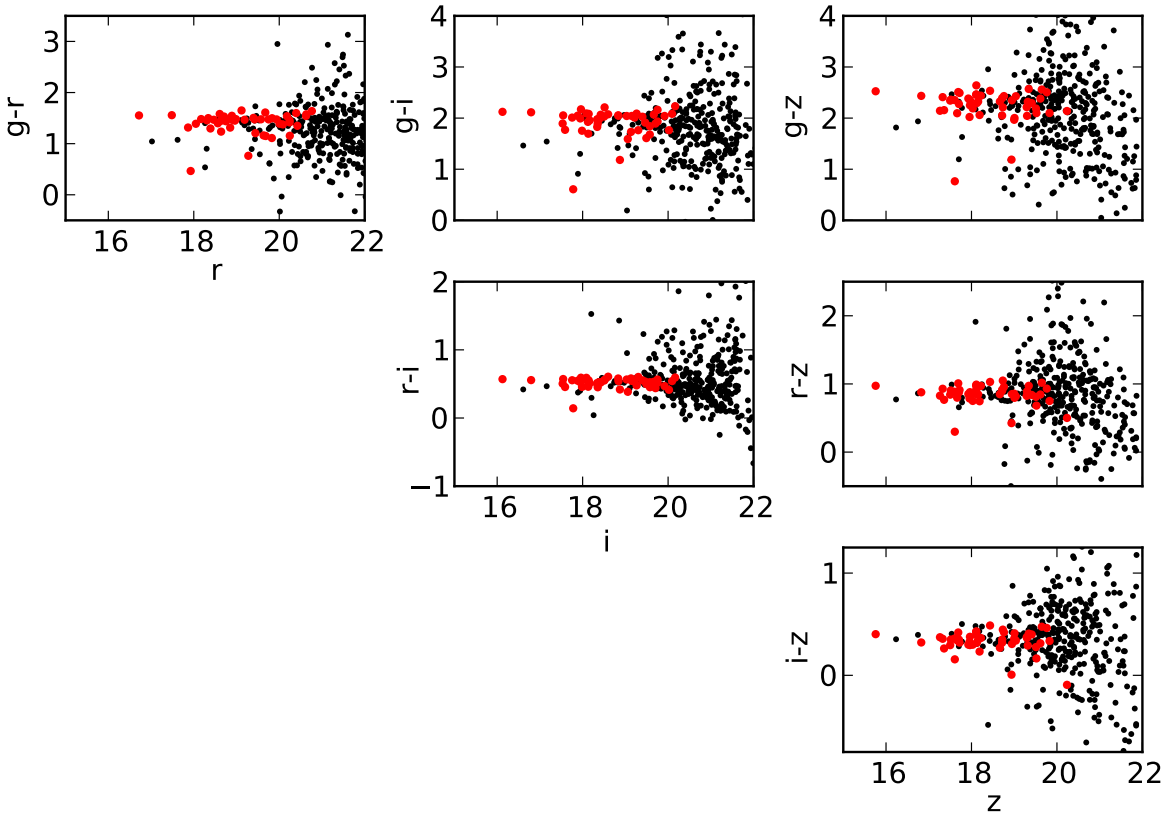

Figure 4.1: Colour-magnitude diagrams for our median mass cluster, Abell0068. Shown are all galaxies within 1Mpc of the cluster centre, with spectroscopically confirmed members marked in red.

Table 4.2: Scaling relation parameters

Bandpass	Intercept a	Slope b	Scatter $\sigma_{\ln L\|M_{\rm WL}}$
	$r_{\rm vir}$		
L_g	$3.16^{+0.39}_{-0.34}$	$1.32^{+0.28}_{-0.42}$	$0.20^{+0.09}_{-0.12}$
L_r	$3.98^{+0.59}_{-0.35}$	$1.36^{+0.29}_{-0.37}$	$0.21^{+0.10}_{-0.12}$
L_i	$5.01^{+0.61}_{-0.55}$	$1.37^{+0.28}_{-0.37}$	$0.20^{+0.09}_{-0.12}$
L_z	$6.17^{+0.91}_{-0.67}$	$1.36^{+0.30}_{-0.39}$	$0.21^{+0.10}_{-0.12}$
L_J	$8.13^{+1.20}_{-0.72}$	$1.37^{+0.28}_{-0.39}$	$0.20^{+0.09}_{-0.12}$
L_K	$14.45^{+1.76}_{-1.27}$	$1.31^{+0.28}_{-0.39}$	$0.21^{+0.10}_{-0.11}$
	r_{500}		
L_g	$3.72^{+0.27}_{-0.25}$	$0.99^{+0.14}_{-0.16}$	$0.10^{+0.05}_{-0.07}$
L_r	$4.79^{+0.34}_{-0.32}$	$1.02^{+0.15}_{-0.17}$	$0.10^{+0.05}_{-0.08}$
L_i	$6.03^{+0.43}_{-0.04}$	$1.01^{+0.13}_{-0.18}$	$0.11^{+0.05}_{-0.08}$
L_z	$7.24^{+0.52}_{-0.48}$	$1.00^{+0.14}_{-0.15}$	$0.11^{+0.05}_{-0.08}$
L_J	$10.00^{+0.72}_{-0.67}$	$1.00^{+0.13}_{-0.16}$	$0.10^{+0.05}_{-0.07}$
L_K	$17.38^{+1.24}_{-1.16}$	$0.97^{+0.13}_{-0.16}$	$0.10^{+0.05}_{-0.07}$
	r_{2500}		
L_g	$3.31^{+0.40}_{-0.29}$	$0.78^{+0.14}_{-0.16}$	$0.12^{+0.06}_{-0.09}$
L_r	$4.57^{+0.56}_{-0.50}$	$0.80^{+0.14}_{-0.17}$	$0.12^{+0.06}_{-0.09}$
L_i	$5.75^{+0.85}_{-0.51}$	$0.78^{+0.15}_{-0.17}$	$0.12^{+0.07}_{-0.09}$
L_z	$6.46^{+0.79}_{-0.57}$	$0.78^{+0.14}_{-0.16}$	$0.12^{+0.06}_{-0.10}$
L_J	$9.12^{+1.11}_{-0.80}$	$0.77^{+0.13}_{-0.15}$	$0.11^{+0.05}_{-0.09}$
L_K	$16.22^{+1.98}_{-1.76}$	$0.77^{+0.15}_{-0.16}$	$0.11^{+0.05}_{-0.08}$

We do not consider selection effects in this work because the effects are diluted as a consequence of our sample being an overlap of several surveys. The LoCuSS "High-L_X" sample was selected on X-ray luminosity, and the overlap with ACReS and SDSS is not dependent on any cluster property. We note that the covariance between $L_{\rm X,RASS}$ and optical/near-infrared luminosity is expected to be minimal and lead to only minor selection effects (Mulroy et al., in prep.).

We perform linear regressions of the total cluster luminosities in 6 bandpasses (*grizJK*) within 3 overdensity radii ($r_{\rm vir}$, r_{500}, r_{2500}) against the weak-lensing cluster masses within the same radii. In Figure 4.2 we show the data points and resultant scaling relation (and 68 per cent confidence region) for each of these bandpass and radius combinations. The scaling relation parameters (intercept, slope, and intrinsic scatter) are shown in Table 4.2, and their trends with wavelength visually presented in Figure 4.3. We note the following features in these results:

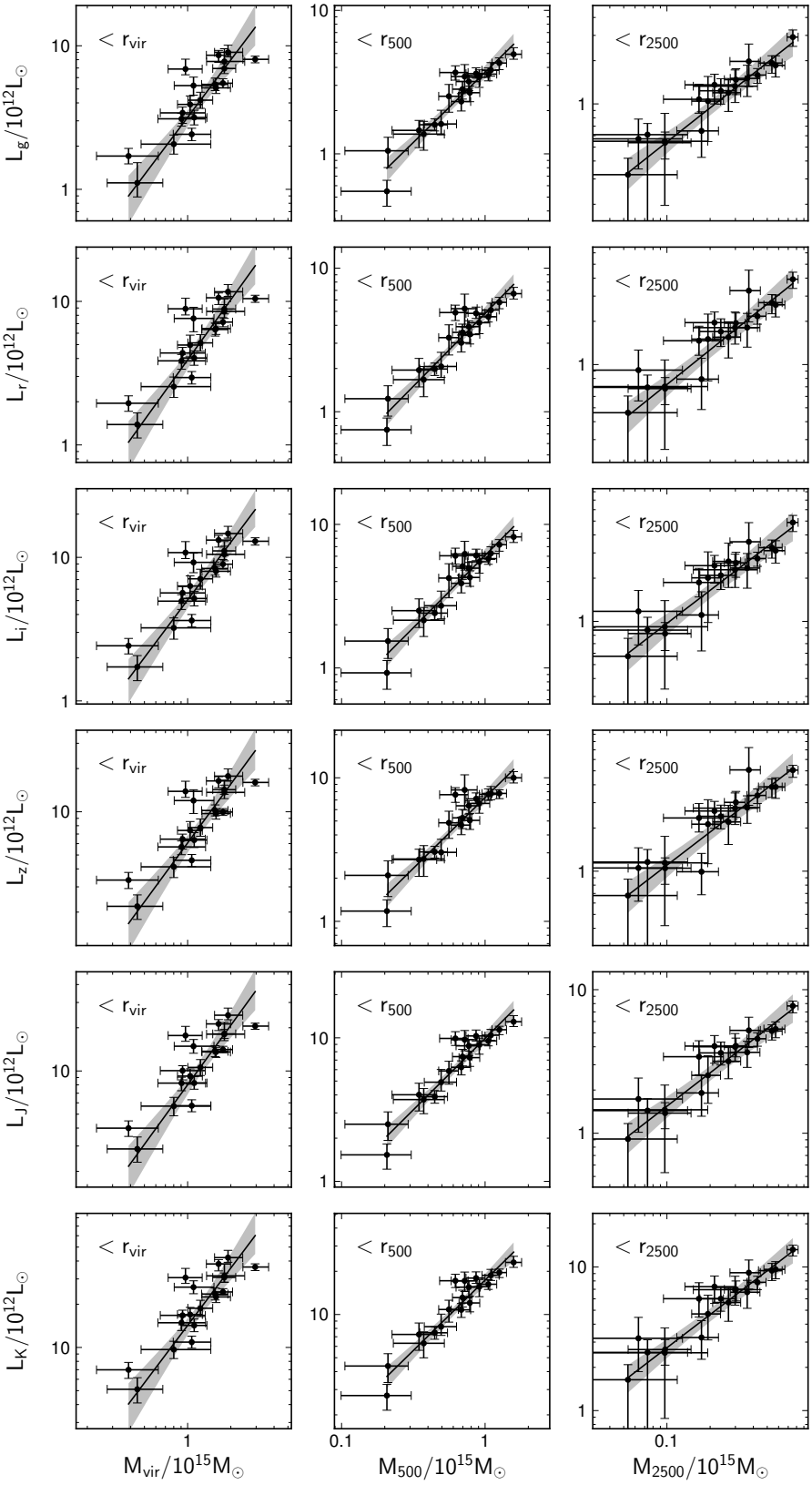

Figure 4.2: Scaling relations between the six total cluster luminosities and weak-lensing cluster mass, where we show the data points, resultant scaling relation and 68 per cent confidence region. Luminosities are calculated from a K band limited sample of galaxies, and both luminosities and masses are measured within $r_{\rm vir}$ [left], r_{500} [middle] and r_{2500} [right].

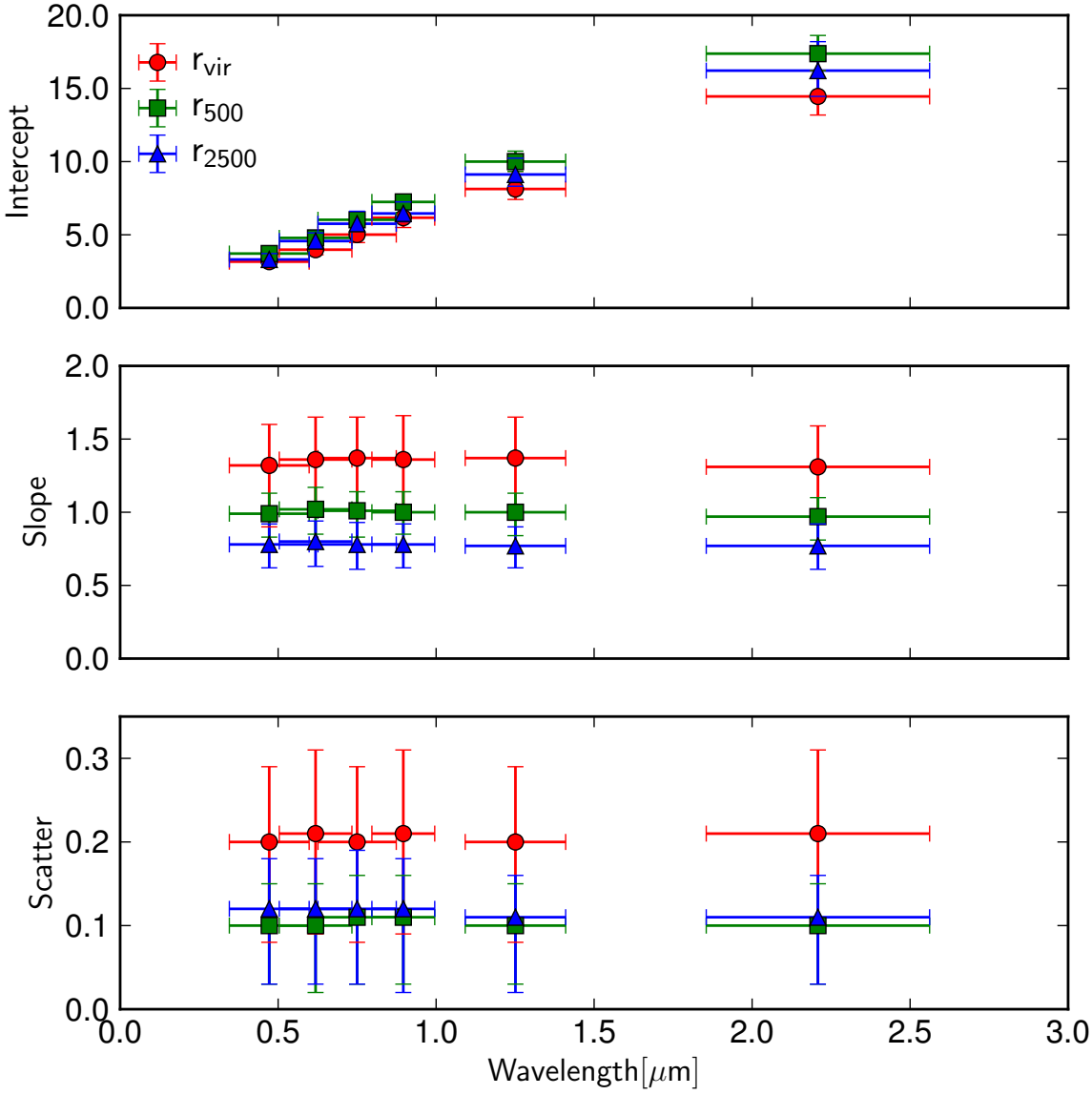

Figure 4.3: Scaling relation parameters (intercept [top], slope [middle] and intrinsic scatter [bottom]) within different radii (red circles: $r_{\rm vir}$, green squares: r_{500}, blue triangles: r_{2500}) as a function of the bandpass wavelength.

1. At fixed radius, both the slope and scatter of the scaling relations are consistent across the wavelength range ($\sim 0.47 - 2.21\mu$m). The same trend is found by Popesso et al. (2005), although our absolute values do not always agree. Some difference is expected due to a different mass measurement method and linear regression scheme. We revisit this observed trend in Section 4.3.2.

2. At fixed radius, the intercept increases with increasing wavelength, as shown in the top panel of Figure 4.3. This rising intercept is a reflection of the SED of the cluster population, which predominantly consists of red galaxies. To understand the shape further we present Figure 4.4, which shows the intercept values at each bandpass for the relations measured within $r_{\rm vir}$, and compare these to values assuming updated Bruzual & Charlot (2003) stellar population models calculated using the tool `EzGal` (Mancone & Gonzalez, 2012) and normalised to the K band of the observations.

   A constant star formation history model (blue lines) for solar metallicity ($Z = 0.02$) and the common assumption of dust extinction ($\tau_v = 1$, Brinchmann et al., 2004; Garn & Best, 2010) is a reasonable match to the observed data in redder bands, but diverges in the bluer bands. Decreasing the dust extinction in such a model ($\tau_v = 0.2$) improves agreement at the red end but increases the discrepancy at the blue end, while increasing the dust extinction ($\tau_v = 5$) improves agreement with the bluest band but is consistently below the observed data. It is not possible to fit the observed data with this model by varying the dust extinction due to the shape of the predictions from a constant star formation history model.

   We find that a single stellar population model (SSP, green lines) and an exponentially decaying model (with a timescale of 1 Gyr, red lines) are almost indistinguishable, and similar in shape to the observed data. These models with solar metallicity agree well with the observed data at the red end, while better agreement across all bands can be found if we allow the metallicity of the galaxies to vary. The observational intercepts of the relations can be well reproduced with either a single stellar population or quickly decaying exponential, with metallicity of ~ 0.4 solar. This is in good agreement with previous studies of the stellar populations of low redshift cluster galaxies (Nelan et al., 2005; Pasquali et al., 2010; Smith et al., 2012).

3. For all relations, the scatter is higher within $r_{\rm vir}$ than within r_{500} or r_{2500}. This is consistent with the increased volume at larger radii allowing for more variation in the large scale structure found within the outskirts of the cluster. As a simple example, infalling groups are more likely to be found in the region between r_{500} and $r_{\rm vir}$, and with greater variation than within r_{500}. Looking at galaxy clusters in the Millennium N-body dark matter simulation (Springel et al., 2005; McGee et al., 2009), we found that at fixed cluster mass, the fractional scatter in the number of galaxies within $r_{\rm vir}$ is roughly double that within both r_{500} and r_{2500}, consistent with our observations.

   A related effect is the fraction of 'interlopers' – spectroscopically confirmed members that are projected onto the cluster but lie beyond the physical radius of the cluster. Given that interlopers are largely uncorrelated with the cluster, a larger interloper fraction suggests a larger variation in that fraction, and therefore a larger inferred scatter. The fraction of members that have not yet passed within r_{200}, quantified using the Millennium simulation as in Haines et al. (2013), is 3.44, 8.76 and 21.14 per cent within projected r_{2500}, r_{500} and $r_{\rm vir}$ respectively, consistent with our observed larger scatter within $r_{\rm vir}$. These interlopers are likely to be bluer than a typical cluster galaxy, and so will slightly increase the intercepts of our relations particularly in the blue bands. This effect will be small, and we don't expect it to affect the results of our analysis in Section 4.3.2 as we work within a single overdensity radius.

4. At fixed wavelength, the slope increases (and consequently the intercept decreases) with increasing radius. This trend is much less prominent when we repeat the analysis without the brightest cluster galaxy (BCG) luminosity, with slopes increasing by ~ 0.05, ~ 0.1 and ~ 0.2 within $r_{\rm vir}$, r_{500} and r_{2500} respectively, resulting in broad agreement in the derived slopes. This suggests that the slope for centrals is shallower and more of a dominant factor at smaller radii. Indeed, this is in agreement with theoretical models that find the stellar mass of a BCG in a cluster of this size does not strongly scale with cluster mass (Behroozi et al., 2013; McCarthy et al., 2017). At larger radii, the luminosity of satellite galaxies (which scales strongly with cluster mass) makes up a larger fraction of the total luminosity and thus drives the slope to be steeper. The remaining trend could be explained if the mass-concentration relation of satellite galaxies is shallower than that of the dark matter.

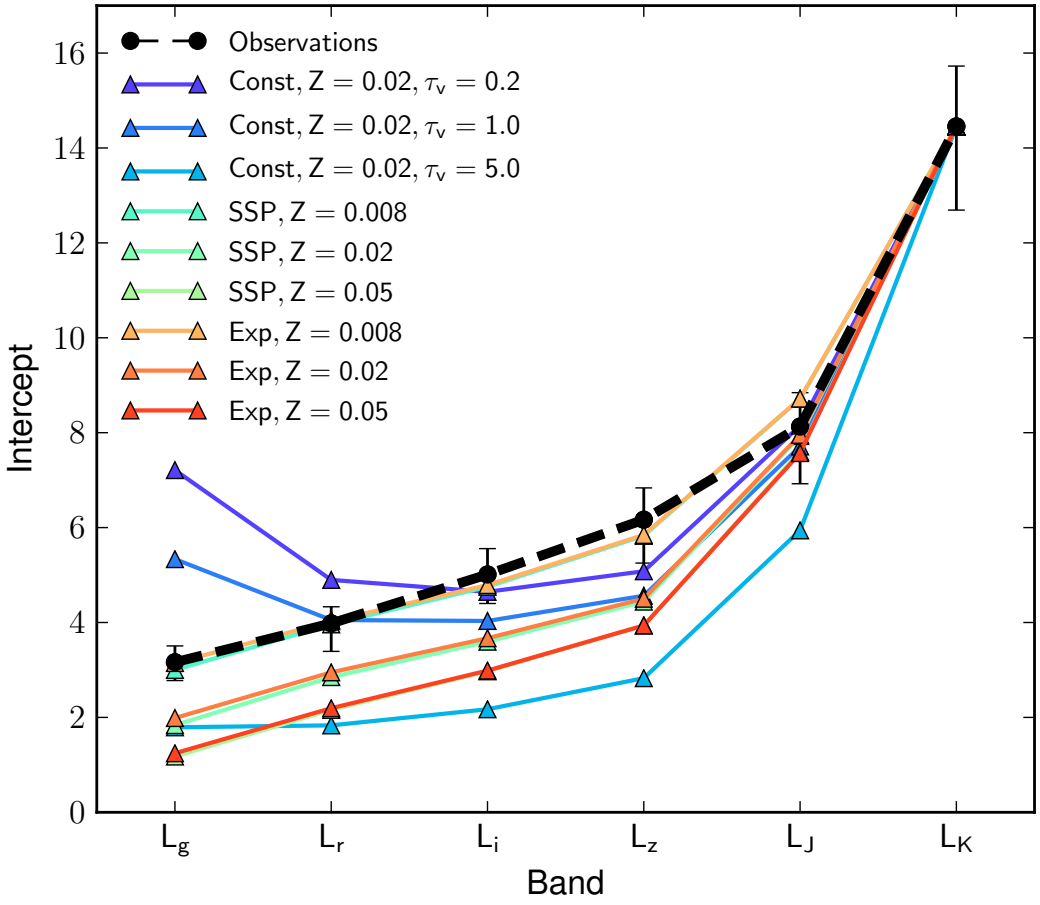

Figure 4.4: Comparison of observed intercepts (black dashed line) with predicted intercept values from updated Bruzual & Charlot (2003) stellar population models (blue lines: constant star formation history model, green lines: single stellar population model, red lines: exponentially decaying model), showing a trend of increasing intercept with increasing wavelength for all models and the observations. All models normalised with respect to the K band value of the observations.

5. The low intrinsic scatter across all wavelengths means that optical/near-infrared light is a good mass proxy for upcoming surveys, as discussed in Section 4.5.

4.3.2 Cluster Colour

In Section 4.3.1, we found that the relation between total luminosity and cluster mass has the same slope across all wavelengths (within a given radius), which suggests that the colour of clusters is not a function of mass. Further, the low scatter in the relations places an upper limit on the variability of the colour of clusters. As shown in Figure 4.5, we find variability in the cluster colours (standard deviation of the distribution of cluster colours) within $r_{\rm vir}$ on the scale

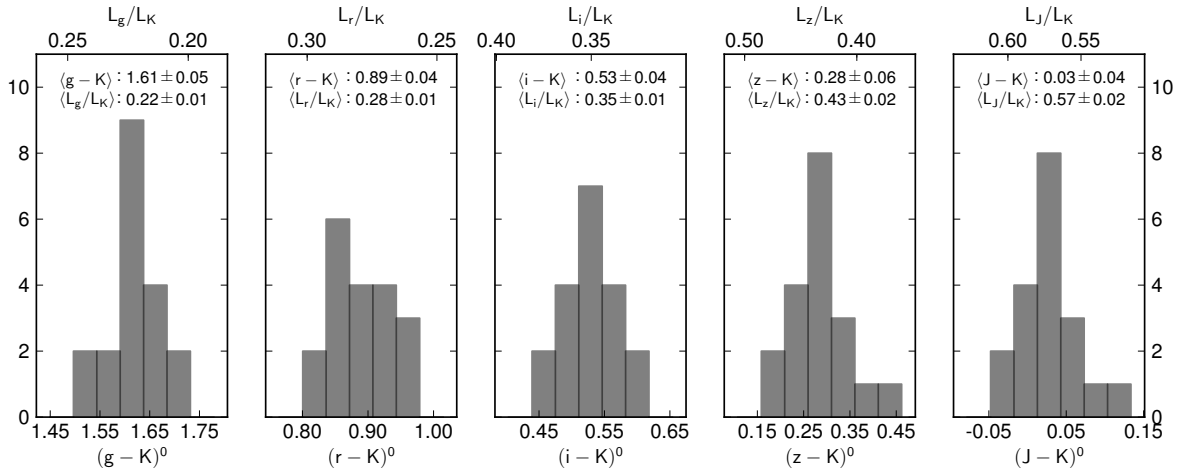

Figure 4.5: Histogram of the rest frame cluster colours within $r_{\rm vir}$.

of $\sigma \sim 0.05$ magnitudes in the full range of colours.

To understand this level of variation, we use updated Bruzual & Charlot (2003) models with a single stellar population, which gave a reasonable match to the intercept of the scaling relations. For an SSP model with a fixed age of 10 Gyr, sampling the galaxy metallicity uniformly in the logarithm between 0.4 solar and 1 solar, leads to an average $(g-K)$ colour variation of $\sigma \sim 0.05$ magnitudes (and similar in other bands). Similarly, at fixed solar metallicity, sampling SSP age between 7 Gyr and 10 Gyr leads to an average colour variation of $\sigma \sim 0.05$ magnitudes. It is worth noting that while this colour variation and the required change in metallicity or age is moderate in terms of individual galaxies, we are considering the mean for each cluster population, for which it is a large variation and requires further investigation to understand.

To investigate the source of this variability we explore the correlation between cluster colour and various indicators of the level of disturbance in that cluster. We consider seven indicators, four of which trace the bulk cluster properties and three of which are driven by the properties of the central region of the cluster.

Bulk Cluster Properties

The DS statistic (Dressler & Shectman, 1988) is a substructure test similar to a χ^2 statistic that quantifies local deviations in mean velocity $\overline{\nu}$ and velocity dispersion σ. For each cluster member the local $\overline{\nu}$ and σ are calculated using $N_{\rm nn}$ nearest neighbours, and compared to the

global cluster values:

$$\delta_i^2 = (N_{\rm nn}/\sigma^2)\left[(\overline{\nu} - \overline{\nu}_{i,\rm local})^2 + (\sigma - \sigma_{i,\rm local})^2\right]. \quad (4.2)$$

The DS statistic, $\Delta_{\rm DS}$, is the sum of δ, and after being normalised by the number of cluster galaxies is ~ 1 for clusters with a Gaussian velocity distribution, with higher values indicating the presence of substructure.

We calculate this statistic considering all members within $r_{\rm vir}$, and using $N_{\rm nn} = \sqrt{N_{\rm members}}$ to keep the measurement consistent between clusters of varying richness, although our measurements are not significantly affected by choosing a fixed number within this range. To quantify the statistical significance of this measurement we also calculate the P-value, by repeating the measurement after random reassignments of member positions to velocities. The P-value is the fraction of times this reassigned measurement is greater than the original DS statistic. We find only three clusters with a P-value > 0.01, corresponding to the three smallest DS statistics, and exclude these values from our analysis.

We define the magnitude gap, $\Delta M_{1,2}$, as the difference in K band magnitude between the two K band brightest cluster members within $0.5{\rm r_{vir}}$. This gives an indication of the time since the last major merger activity in a cluster; a smaller gap suggests more recent infall of bright galaxies, while a larger gap suggests that the bright central galaxies have had time since any significant merger event to accrete onto the BCG (e.g. Dariush et al., 2010; Deason et al., 2013).

We also calculate the projected separation between the X-ray centroid (Martino et al., 2014) and the BCG position, $\Delta^{\rm BCG}_{\rm X\text{-}ray}$. In a dynamically relaxed cluster both the X-ray emitting hot gas and the BCG are centred on the minimum of the gravitational potential well, and so a larger separation indicates a more disturbed cluster.

Finally, we use the centroid shift parameter, $\langle w \rangle$, calculated in Martino et al. (2014) as a measure of the cluster X-ray morphology. It is defined as the standard deviation of the projected separation between the X-ray peak and the X-ray centroid calculated in circular apertures in the range $[0.05 - 1]r_{500}$. Clusters with high centroid shift are typically disturbed clusters, while those with low centroid shift are typically more relaxed. We note that as both $\Delta^{\rm BCG}_{\rm X\text{-}ray}$ and $\langle w \rangle$ are projected separations, they are insensitive to separation along the line of sight.

Central Cluster Properties

In the centre of some clusters, the ICM is strongly radiating and cooling. The cores of these clusters are therefore cool and dense, with low entropy and high surface brightness (e.g. Poole et al., 2008; Rossetti & Molendi, 2010). We use three parameters to probe the presence of these cool cores, and therefore to indicate the dynamical state of the ICM.

Following Santos et al. (2008) we calculate the surface brightness concentration, c_{SB}, as the ratio of the peak central surface brightness and the ambient surface brightness:

$$c_{SB} = \frac{\mathrm{SB}(< 40\mathrm{kpc})}{\mathrm{SB}(< 400\mathrm{kpc})}. \tag{4.3}$$

The surface brightness probes the emission of the ICM, and so a higher surface brightness concentration suggests the presence of a cool core, and therefore a more relaxed system.

Sanderson et al. (2009) calculate α, the logarithmic slope of the gas density profile at $0.04r_{500}$ ($\sim$ 40kpc for these objects) for all but one (Abell0291) of the clusters in our sample. The gas density slope traces the temperature slope, which steepens with increased cooling, and so a more negative α implies stronger cooling.

Also from Sanderson et al. (2009), we use central entropy, K, measured within 20kpc and defined as $K = Tn_e^{-2/3}$, where T is the cluster temperature and n_e is the electron density. As a measure of the thermal history of the ICM, lower entropy is associated with the presence of a cool core.

Cluster Colour Trends

In Figure 4.6 we show the total rest frame $(g - K)$ cluster colour within r_{vir} as a function of the seven indicators of disturbance discussed above. The bulk cluster properties (Δ_{DS}, $\Delta M_{1,2}$, $\Delta^{\mathrm{BCG}}_{\mathrm{X\text{-}ray}}$ and $\langle w \rangle$) all show a trend of decreasing scatter ($\sigma_{(g-K)}$) in cluster colour as clusters become more disturbed. c_{SB} also suggests this trend, however it does not appear in the other central cluster properties (α and K). Note that we show the values for $(g - K)$ colour within r_{vir}, because it covers the widest wavelength baseline and the whole cluster, but the results are similar with other choices of colour and radii. As an example, we show in Figure 4.7 the trend of cluster colour with the DS statistic, where the colour is defined over a range of colours ($g - K$, $r - K$, $i - K$, $z - K$, $J - K$) and calculated within a range of radii (r_{vir}, r_{500}, r_{2500}). The

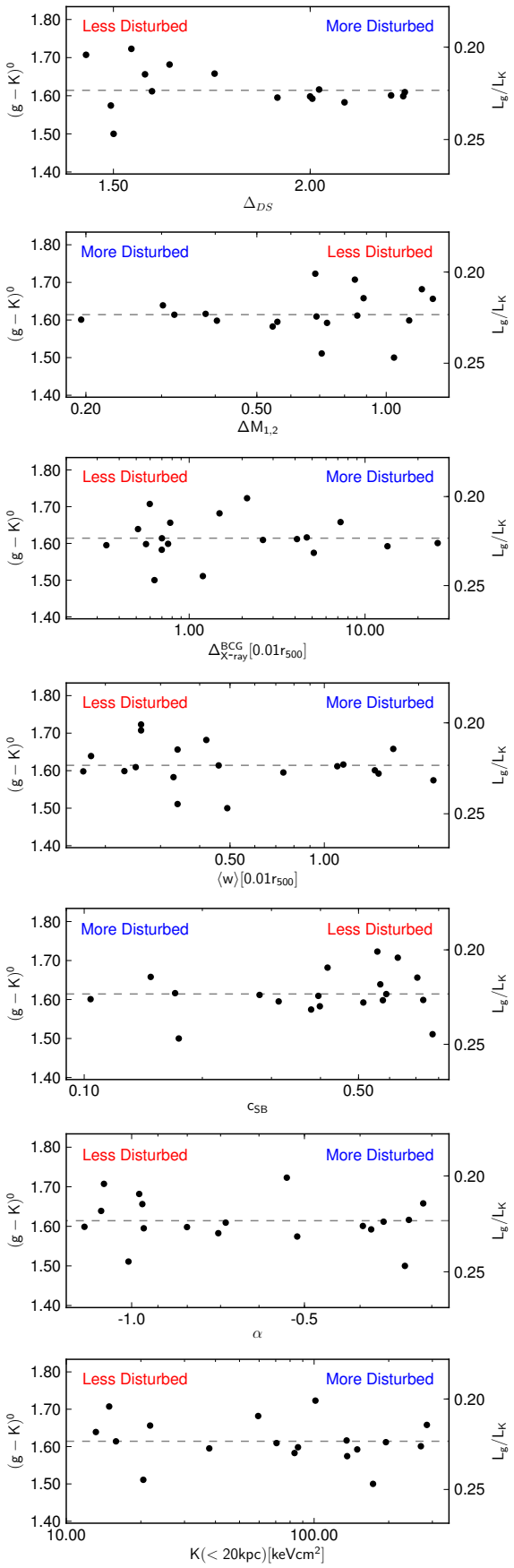

Figure 4.6: Rest frame $(g-K)$ cluster colour within $r_{\rm vir}$ as a function of various indicators of the level of disturbance in the cluster. From top to bottom: DS statistic; magnitude gap; BCG / X-ray centroid separation; centroid shift; surface brightness concentration; alpha, the logarithmic slope of the gas density; and central entropy.

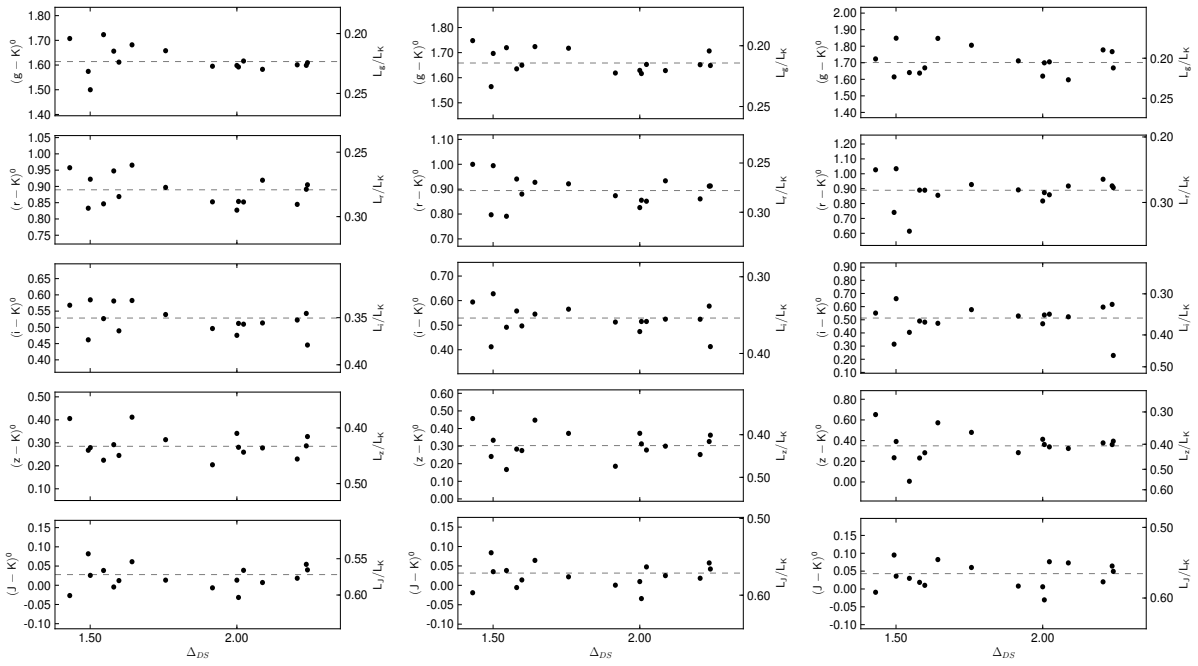

Figure 4.7: Rest frame $(g-K)$ cluster colour within $r_{\rm vir}$ [left], r_{500} [middle] and r_{2500} [right] as a function of the DS statistic.

trend seen in the top left panel and discussed here is also visible in most other colour/radius combinations.

To quantify this trend, we split the full sample based on each indicator into two subsamples (disturbed and undisturbed), either splitting the sample in half or where there appears to be a natural division near the median. We then calculate the spread in the cluster colour within these subsamples. As shown in Table 4.3 and Figure 4.8 there is a clear difference between the two subsamples for most indicators, with the disturbed clusters showing less variability in cluster colour than the undisturbed clusters. Most interestingly, the degree of variation in the subsamples varies systematically with the cluster disturbance indicator by which the sample was split. The properties towards the left of Figure 4.8 are the bulk properties, thereby indicating disturbance on large scales, while those towards the right are the ICM properties and as such probe closer to the cluster centre. For instance, the farthest left parameter (Δ_{DS}) measures disturbances on the scale of the whole cluster, and is often used to detect infalling galaxy groups. Similarly, the $\Delta M_{1,2}$ parameter measures disturbances within $0.5r_{\rm vir}$ (the region for which a second-rank galaxy is searched). The disturbance indicators probe smaller and smaller scales, until the right-most indicator on the figure, which probes the cluster central entropy

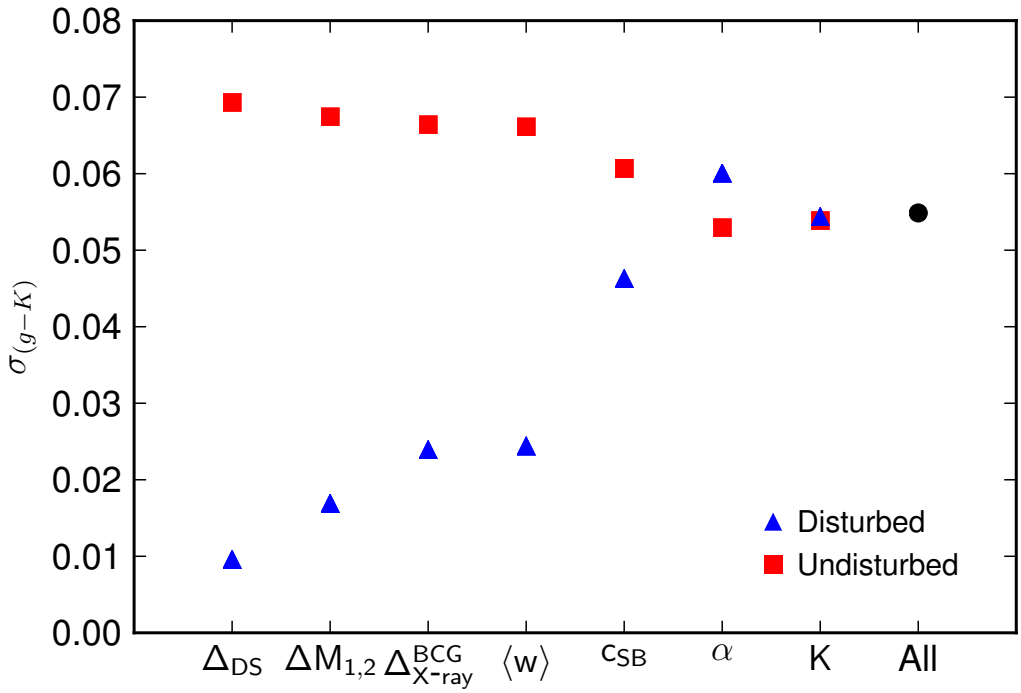

Figure 4.8: Variation in cluster colour within r_{vir}, $\sigma_{(g-K)}$, for two subsamples (blue triangles: disturbed, red squares: undisturbed) defined by various indicators of the level of disturbance in the cluster.

within the central 20kpc. Taken together, these results suggest that there is a larger spread in the stellar age, metallicity and/or SFR in undisturbed clusters than in disturbed clusters, and that this effect decreases with disturbance indicators towards the cluster centre.

4.4 Interpretation

There are three broad potential causes of variation in cluster colour. We will discuss each of these in turn:

1. The role of the BCG - the state of the cluster (disturbed or undisturbed) can strongly affect the colour of the BCG. It is known that undisturbed cool core clusters have BCGs with a greater range of SFRs and optical emission lines (Cavagnolo et al., 2008; McDonald et al., 2010). If the BCG colour dominated the total cluster colour, then this would lead to an increase in the variability of the cluster colour in undisturbed clusters with strongly cooling cores. While the observed trend is in the same direction, we would expect to see the biggest trend in indicators which probe near the cluster core (α, K), but we see no

Table 4.3: Variation within subsamples

Indicator	Disturbed Clusters	Undisturbed Clusters	Indicator Threshold
	$\langle g-K\rangle \pm \sigma_{(g-K)}$		
All	1.614 ± 0.055		
Δ_{DS}	1.599 ± 0.010	1.639 ± 0.069	1.8
$\Delta M_{1,2}$	1.607 ± 0.017	1.619 ± 0.067	0.6
$\Delta^{BCG}_{X\text{-ray}}$	1.609 ± 0.024	1.617 ± 0.066	2.5
$\langle w\rangle$	1.607 ± 0.024	1.618 ± 0.066	0.6
c_{SB}	1.603 ± 0.046	1.627 ± 0.061	0.5
α	1.610 ± 0.060	1.618 ± 0.053	-0.6
K	1.606 ± 0.052	1.625 ± 0.057	75.0
	$\langle L_g/L_K\rangle \pm \sigma_{(L_g/L_K)}$		
All	0.222 ± 0.011		
Δ_{DS}	0.225 ± 0.002	0.217 ± 0.014	1.8
$\Delta M_{1,2}$	0.224 ± 0.003	0.222 ± 0.014	0.6
$\Delta^{BCG}_{X\text{-ray}}$	0.223 ± 0.005	0.222 ± 0.014	2.5
$\langle w\rangle$	0.224 ± 0.005	0.222 ± 0.014	0.6
c_{SB}	0.224 ± 0.010	0.220 ± 0.012	0.5
α	0.223 ± 0.012	0.221 ± 0.011	-0.6
K	0.224 ± 0.011	0.220 ± 0.012	75.0

such trend in these indicators. Additionally, the BCG luminosity is typically only $\sim 5\%$ of the total cluster luminosity, so is subdominant.

2. Infalling galaxies - it is well known that galaxies within massive halos, such as galaxy groups and clusters, have systematically less star formation than isolated galaxies (e.g. McGee et al., 2011; Wetzel et al., 2012). Furthermore, the fraction of star forming galaxies is remarkably similar in different groups (Balogh & McGee, 2010). Therefore, it could be the case that while undisturbed clusters are continually accreting star forming field galaxies which are quenched as they fall into the cluster, disturbed clusters are gaining their mass from infalling groups and clusters. The galaxies in these groups already reside in a dense environment, and so have already had their star formation quenched. As a result these galaxies have little impact on the overall star formation of the cluster, in contrast to the field galaxies falling into the undisturbed clusters and introducing cluster to cluster variation. This effect would decrease towards the cluster centre, consistent with the trend being clear in the bulk cluster properties but only in one of the three centre

cluster properties.

We have tested this hypothesis using accretion histories for clusters from the Millennium N-body dark matter simulation (Springel et al., 2005; McGee et al., 2009). However, we find that clusters which are currently undergoing a major merger have not accreted a significantly higher fraction of their galaxies through massive halos in the last 1 - 4 Gyrs. For instance, the fraction of galaxies accreted through haloes of mass $> 10^{13}$ $M_{\odot}$ in the last 2 Gyrs in clusters undergoing a major merger is 0.35 $\pm$ 0.01, while it is 0.34 $\pm$ 0.01 in clusters not undergoing a major merger.

3. The effect of mergers - there has been recent evidence that major mergers may affect the star formation properties of the galaxies within them (Rawle et al., 2014; Pranger et al., 2014; Stroe et al., 2015). If a merger could 'standardise' the SFR in a cluster, then disturbed clusters would have less variation in their total colour. As time passes since the last major merger, the spread in SFR and therefore cluster colour would increase.

   The merger could standardise the SFR by leading to a burst and/or quenching, as long as it led to a similar effect in all merging clusters. One possible scenario is galaxy interaction with the shocks created by merging clusters. The Mach number of a galaxy is typically $\mathcal{M} \sim 1$ (Sarazin, 1988), while that of a cluster shock can be as high as 5 (e.g. van Weeren et al., 2010), so we would expect to see the effects of a shock across the entire cluster well before the galaxies were virialised. The standardisation would be seen most clearly in the cluster disturbance probes which examine the widest range (e.g. Δ_{DS}), in good agreement with the observed trends we see. These cluster-wide shocks are unlikely to alter the densest gas in the cluster core, and so would not be detectable in the cluster disturbance indicators which probe the central ICM properties (e.g. K) (Poole et al., 2008). Any disruption to the cool core would occur during the later stages of a merger.

Given all this, it seems that the merger itself, and perhaps the shock it triggers, is the most likely cause of the lack of variation we see in the total cluster colour of disturbed clusters. While the precise physical mechanism which causes this is unclear, upcoming low frequency radio facilities (eg., LOFAR; van Haarlem et al., 2013) will find hundreds of merging clusters whose shock waves can be mapped by their radio emission, and should lead to tighter constraints on

the physical mechanism.

4.5 Conclusions and Implications for Future Surveys

In this study, we have used measurements of the luminosities and colours of 19 galaxy clusters with well measured weak-lensing masses, highly complete stellar mass limited spectroscopy, and a wide range of indicators of the levels of disturbance in the clusters. We can summarise our main conclusions as:

1. The slope and scatter of the relation between total cluster luminosity and cluster mass is consistent across the full range of bandpasses we probed ($grizJK$). The trend in intercept of these relations is well understood if the galaxy clusters are made up of predominantly old, passive galaxies with metallicities ~ 0.4 solar.

2. The intrinsic scatter in these relations is ~ 0.1 within r_{500} and suggests they would be good, cheap mass proxies for large scale photometric surveys of galaxy clusters, as discussed further below.

3. The variation in cluster colour shows trends with the overall cluster disturbance, increasing as clusters become more relaxed, perhaps indicating that the major mergers are a standardising force in the global colours, possibly through system-wide shocks.

We have shown that total cluster luminosity scales tightly with weak-lensing mass over the full range of wavelengths considered here (Table 4.2, Figures 4.2 & 4.3). Combined with the fact that these measurements were made on shallow survey data, this work suggests that these luminosities are promising mass proxies for future surveys, consistent with previous studies (e.g. Girardi et al., 2000; Lin et al., 2003, 2004; Ramella et al., 2004; Popesso et al., 2005; Mulroy et al., 2014; Pearson et al., 2015; Ziparo et al., 2016). We highlight that the specific luminosity measurements used in this work benefited from highly complete K band limited (roughly stellar mass limited) spectroscopic membership catalogues, and prior radial knowledge from weak-lensing analysis. Future studies will be needed to quantify the best method for luminosity measurements in the absence of this prior information.

The observed wavelength range considered ($\sim 0.47 - 2.21\mu$m) corresponds to a rest frame wavelength range of $\sim 0.38 - 1.80\mu$m at our average redshift $\langle z \rangle = 0.23$, for which we have

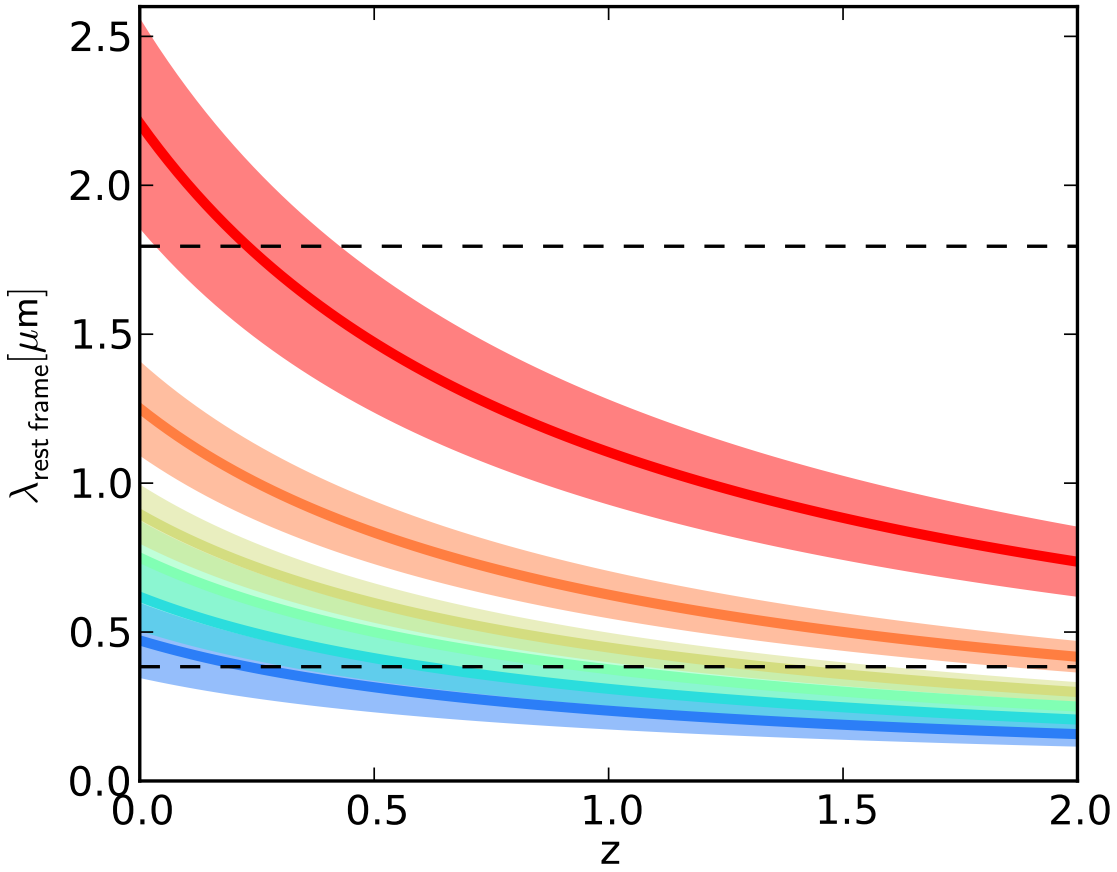

Figure 4.9: Rest frame wavelength for each bandpass (*grizJK*, shown as blue through red) as a function of redshift.

shown these luminosities to tightly scale with mass. The redshift evolution of this rest frame wavelength can be seen in Figure 4.9, which highlights the importance of observed near-infrared wavelengths when studying clusters at redshifts of 1 and above, which is significant for ongoing and upcoming surveys and instruments such as DES, HSC, Euclid, and LSST.

Chapter 5

Galaxy Cluster Scaling Relations

This chapter is taken from Mulroy et al. in preparation: "LoCuSS: Galaxy Cluster Scaling Relations", and involves the use of a Bayesian code developed by a co-author and X-ray & SZ parameters calculated by co-authors.

5.1 Introduction

In Chapter 1 we demonstrated that the mass of galaxy clusters is an important probe of cosmology. In Chapter 2 we introduced scaling relations as a way to determine these masses, and highlighted that the selection function of the sample used to parameterise these relations can bias the results. While there has been a lot of work on galaxy cluster scaling relations, only a small number take these selection effects into account, and those that do are limited to X-ray observables (Vikhlinin et al., 2009; Pratt et al., 2009; Mantz et al., 2016b).

A multivariate approach to linear regression is able to correct for selection effects by considering the selection variable alongside other cluster observables and constraining the covariance matrix between observables at fixed mass. These constraints on covariance also provide an insight into how different astrophysical processes interact within clusters. To date only a handful of observational constraints have been placed on these values, and as with corrected scaling relations, are limited to X-ray observables (Mantz et al., 2010; Maughan, 2014; Mantz et al., 2015, 2016a,b).

In this chapter we introduce a new maximum likelihood method to parameterise the scaling relations and covariances using a sample of 42 galaxy clusters. This sample has well constrained observables measured from a broad range of instruments (*Chandra*/*XMM-Newton*, Subaru, Hectospec, UKIRT and Planck), making this the first study to constrain these parameters across

the electromagnetic spectrum.

We introduce our sample and data in Section 5.2, and our maximum likelihood method in Section 5.3. Our scaling relation and covariance results are presented in Section 5.4, and we summarise in Section 5.5. We assume $\Omega_{\rm m,0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$ and $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$. In this cosmology, at the average cluster redshift, $\langle z \rangle = 0.22$, 1 arcsec corresponds to a projected physical scale of 3.55 kpc.

5.2 Data and Analysis

5.2.1 Sample

We study a sample of 42 X-ray luminous clusters from the "High-L_X" sample of the Local Cluster Substructure Survey (LoCuSS[1]), which was selected from the ROSAT All Sky Survey (RASS, Ebeling et al. 1998, 2000; Böhringer et al. 2004). These are all the clusters satisfying a clearly defined selection criteria: $n_H < 7\times10^{20}{\rm cm}^{-2}$; $-25^\circ < \delta < +65^\circ$; $L_{X,RASS}E(z)^{-1} > 4.4\times10^{44}{\rm erg/s}$ for clusters between $0.15 < z < 0.24$, and $L_{X,RASS}E(z)^{-1} > 7.0\times10^{44}{\rm erg/s}$ for clusters between $0.24 < z < 0.30$ (Table 5.1, Figure 5.1), where $E(z) \equiv H(z)/H_0 = \sqrt{\Omega_{\rm m,0}(1+z)^3 + \Omega_{\Lambda,0}}$ is the evolution of the Hubble parameter. All observables described below are summarised in Table 5.2.

RASS L_X

The $L_{X,{\rm RASS}}$ values cover the soft X-ray band from 0.1 to 2.4 keV, and are taken from the ROSAT Brightest Cluster Sample and its low flux extension (BCS, Ebeling et al. 1998; eBCS, Ebeling et al. 2000) for objects in the northern hemisphere, and the ROSAT-ESO Flux Limited X-ray galaxy cluster survey (REFLEX, Böhringer et al., 2004) for objects mostly in the southern hemisphere ($\delta < 2.5^\circ$). For the clusters in the overlap between surveys (Abell0267: BCS, REFLEX and Abell2631: eBCS, REFLEX) we average the luminosities. RASS luminosities cannot be core-excised due to the RASS PSF, and so are sensitive to the presence or lack thereof of a cool core, which will increase the luminosity.

[1] http://www.sr.bham.ac.uk/locuss

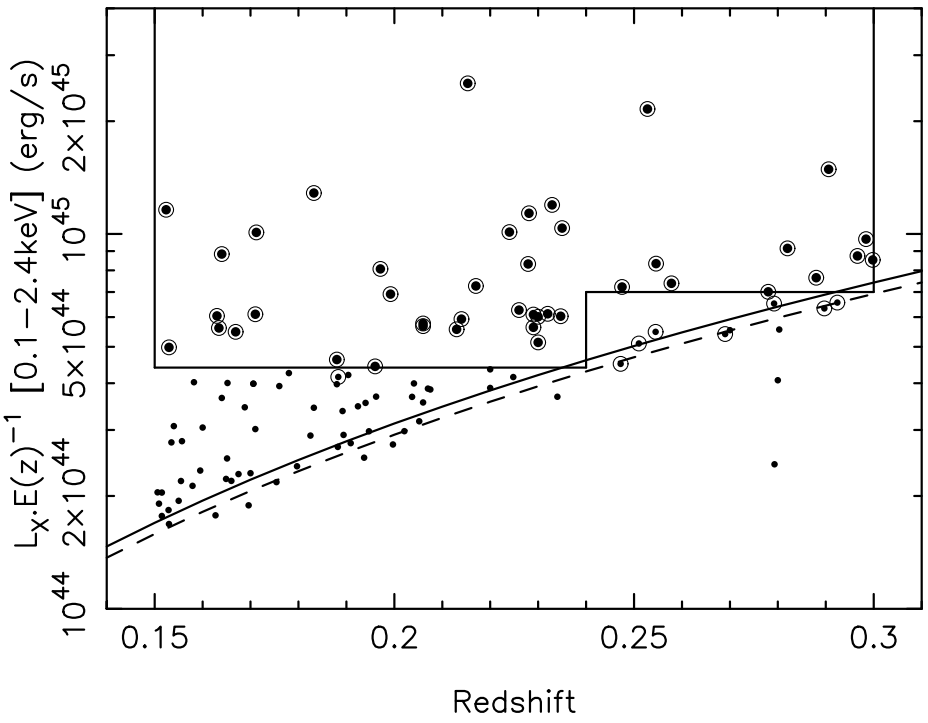

Figure 5.1: The $L_{X,\mathrm{RASS}}E(z)^{-1}-$ redshift distribution of the LoCuSS clusters; The large points show the 42 clusters passing the selection criteria and therefore used in this work, while the circles show the 50 LoCuSS "High-L_X" clusters. The straight lines represent the selection criteria, the dashed curve is the 75% completeness limit for (e)BCS (Ebeling et al., 1998, 2000) and the solid curve is the 90% completeness limit for REFLEX (Böhringer et al., 2004).

Table 5.1: Cluster sample

Name	RA α [J2000]	Dec δ [J2000]	Redshift z	$L_{X,\mathrm{RASS}}$ 10^{44}erg/s	M_{WL} $10^{14}M_{\odot}$
Abell2697	0.7990	-6.0860	0.2320	$6.88^{+0.85}_{-0.85}$	$6.61^{+1.20}_{-1.21}$
Abell0068	9.2785	9.1566	0.2546	$9.47^{+2.61}_{-2.61}$	$6.82^{+1.11}_{-1.01}$
Abell0115	14.0012	26.3424	0.1971	$8.90^{+2.13}_{-2.13}$	$5.39^{+1.62}_{-1.49}$
Abell0141	16.3864	-24.6466	0.2300	$5.76^{+0.90}_{-0.90}$	$4.56^{+0.92}_{-0.86}$
Abell0209	22.9689	-13.6112	0.2060	$6.29^{+0.65}_{-0.65}$	$12.34^{+1.64}_{-1.50}$
Abell0267	28.1748	1.0072	0.2300	$6.74^{+1.42}_{-1.42}$	$5.60^{+0.91}_{-0.85}$
Abell0291	30.4296	-2.1966	0.1960	$4.88^{+0.56}_{-0.56}$	$4.46^{+1.02}_{-0.95}$
Abell0521	73.5287	-10.2235	0.2475	$8.18^{+1.36}_{-1.36}$	$5.39^{+0.99}_{-0.93}$
Abell0586	113.0845	31.6335	0.1710	$6.64^{+1.30}_{-1.30}$	$7.21^{+1.60}_{-1.40}$
Abell0611	120.2367	36.0566	0.2880	$8.86^{+2.53}_{-2.53}$	$9.11^{+1.67}_{-1.56}$
Abell0697	130.7398	36.3666	0.2820	$10.57^{+3.28}_{-3.28}$	$7.71^{+1.54}_{-1.43}$
ZwCl0857.9+2107	135.1536	20.8946	0.2347	$6.79^{+1.76}_{-1.76}$	$2.07^{+0.99}_{-1.08}$
Abell0750	137.3024	10.9745	0.1630	$6.59^{+1.40}_{-1.40}$	$6.15^{+1.71}_{-1.35}$
Abell0773	139.4726	51.7271	0.2170	$8.10^{+1.35}_{-1.35}$	$10.07^{+1.07}_{-1.00}$
Abell0781	140.1075	30.4941	0.2984	$11.29^{+2.82}_{-2.82}$	$4.75^{+1.72}_{-1.89}$
ZwCl0949.6+5207	148.2048	51.8849	0.2140	$6.60^{+1.15}_{-1.15}$	$4.97^{+1.13}_{-1.04}$
Abell0901	149.1099	-9.9560	0.1634	$6.08^{+0.58}_{-0.58}$	$2.79^{+0.81}_{-0.71}$
Abell0907	149.5917	-11.0640	0.1669	$5.95^{+0.49}_{-0.49}$	$11.52^{+1.95}_{-1.67}$
Abell0963	154.2652	39.0471	0.2050	$6.39^{+1.19}_{-1.19}$	$6.96^{+1.11}_{-1.03}$
ZwCl1021.0+0426	155.9152	4.1863	0.2906	$17.26^{+2.93}_{-2.93}$	$5.32^{+0.87}_{-0.82}$
Abell1423	179.3223	33.6110	0.2130	$6.19^{+1.34}_{-1.34}$	$4.44^{+0.89}_{-0.81}$
Abell1451	180.8199	-21.5484	0.1992	$7.63^{+1.63}_{-1.63}$	$8.17^{+1.04}_{-0.96}$
ZwCl1231.4+1007	188.5728	9.7662	0.2290	$6.32^{+1.58}_{-1.58}$	$4.61^{+1.44}_{-1.47}$
Abell1682	196.7083	46.5593	0.2260	$7.02^{+1.37}_{-1.37}$	$8.52^{+1.06}_{-0.99}$
Abell1689	197.8730	-1.3410	0.1832	$14.07^{+1.13}_{-1.13}$	$12.57^{+1.53}_{-1.40}$
Abell1763	203.8337	41.0012	0.2279	$9.32^{+1.33}_{-1.33}$	$15.80^{+2.16}_{-1.94}$
Abell1835	210.2588	2.8786	0.2528	$24.48^{+3.35}_{-3.35}$	$10.97^{+1.56}_{-1.44}$
Abell1914	216.4860	37.8165	0.1712	$10.98^{+1.11}_{-1.11}$	$7.83^{+1.35}_{-1.24}$
ZwCl1454.8+2233	224.3131	22.3428	0.2578	$8.41^{+2.10}_{-2.10}$	$3.74^{+1.46}_{-1.44}$
Abell2009	225.0813	21.3694	0.1530	$5.37^{+0.99}_{-0.99}$	$6.39^{+1.45}_{-1.25}$
RXCJ1504.1-0248	226.0313	-2.8047	0.2153	$28.07^{+1.49}_{-1.49}$	$6.54^{+1.48}_{-1.32}$
Abell2111	234.9188	34.4243	0.2290	$6.83^{+1.65}_{-1.65}$	$5.09^{+1.39}_{-1.21}$
Abell2204	248.1956	5.5758	0.1524	$12.50^{+1.34}_{-1.34}$	$9.92^{+1.82}_{-1.59}$
Abell2219	250.0827	46.7114	0.2281	$12.73^{+1.37}_{-1.37}$	$8.65^{+1.34}_{-1.29}$
RXJ1720.1+2638	260.0420	26.6257	0.1640	$9.57^{+1.07}_{-1.07}$	$4.94^{+1.38}_{-1.17}$
Abell2261	260.6133	32.1326	0.2240	$11.31^{+1.55}_{-1.55}$	$10.75^{+1.30}_{-1.20}$
RXCJ2102.1-2431	315.5411	-24.5335	0.1880	$5.07^{+0.55}_{-0.55}$	$3.71^{+0.87}_{-0.79}$
RXJ2129.6+0005	322.4165	0.0894	0.2350	$11.66^{+2.92}_{-2.92}$	$3.46^{+1.14}_{-1.22}$
Abell2390	328.4034	17.6955	0.2329	$13.43^{+3.14}_{-3.14}$	$10.53^{+1.52}_{-1.41}$
Abell2537	347.0926	-2.1921	0.2966	$10.17^{+1.45}_{-1.45}$	$8.57^{+2.03}_{-1.82}$
Abell2552	347.8887	3.6349	0.2998	$9.94^{+2.84}_{-2.84}$	$7.16^{+1.88}_{-1.69}$
Abell2631	354.4155	0.2714	0.2779	$8.07^{+2.11}_{-2.11}$	$5.61^{+1.58}_{-1.78}$

5.2.2 Gravitational Weak-Lensing Masses

We use weak-lensing masses from Okabe & Smith (2016), who calculate masses by fitting an NFW (Navarro, Frenk & White, 1997) mass profile to the shear profile obtained from Subaru/Suprime-Cam observations. We use M_{500} values, defined as the mass within r_{500}, the radius within which the average density is $500 \times \rho_{\rm crit}$, where $\rho_{\rm crit} = 3H(z)^2/8\pi G$, the critical density of the Universe. We adopt these weak-lensing determined radii, $r_{500,\rm WL}$, as the radii within which we measure most other observables for this work.

5.2.3 X-Ray Data

We use X-ray measurements of the ICM as calculated in Martino et al. (2014), where the full "High-L_X" sample has been observed with either or both of *Chandra* and *XMM-Newton*, and measurements of those in the overlap agree within a few per cent between satellites. Bolometric L_X and T_X are measured within an annulus of $[0.15-1]r_{500,\rm WL}$ to avoid the measurements being dominated by emission from the core, and $M_{\rm gas}$ is measured within $r_{500,\rm WL}$.

We also measure Y_X, proposed by Kravtsov et al. (2006) as the X-ray equivalent of the Sunyaev-Zel'dovich effect described in Section 5.2.4. It is proportional to the product of $M_{\rm gas}$ and T_X, and as such is a measure of the total thermal energy of the ICM. It is measured within its own iteratively defined r_{500}, and cannot be measured for the two clusters with only *Chandra* ACIS-S data (Abell0611 and ZwCl0949.6+5207).

5.2.4 Sunyaev-Zel'dovich Effect

The Sunyaev-Zel'dovich effect is caused by the inverse compton scattering of CMB photons by hot electrons, in this case in the ICM. These interactions boost the photon energy by $\sim k_B T/m_e c^2$, increasing the signal at $\gtrsim$ 218GHz. Multiplied by the optical depth and summed over all electrons in the ICM, this is a direct measure of the thermal energy of the ICM.

We calculate $Y_{\rm Pl}$ ($Y_{\rm SZ}$) using signal maps from the Planck High Frequency Instrument (Planck Collaboration et al., 2016a). The $Y_{\rm Pl}$ flux is computed within $5r_{500,\rm WL}$ using an Arnaud et al. (2010) pressure template. While we use the known cluster positions, the Planck team identify clusters as peaks in the signal map with a signal to noise above 4, and as such identify 39 of

the 42 clusters in our sample. For this overlap our flux measurements are in agreement with those measured by the Matched Multi-Filter 3 (MMF3) algorithm (Planck Collaboration et al., 2016c).

5.2.5 Optical Observables

Near-Infrared Luminosity

To investigate the stellar content of the clusters, we use near-infrared data (Haines et al., 2009), where 39 were observed with WFCAM on UKIRT, and two (Abell0963 and ZwCl0857) with NEWFIRM on the Mayall 4-m telescope at Kitt Peak National Observatory. We lack near-infrared data for Abell2697.

We analyse the data similar to Chapter 3 (Mulroy et al., 2014), selecting cluster members as galaxies lying along a ridge line in $(J-K)/K$ space. We select those within $r_{500,\mathrm{WL}}$ of the cluster centre down to a magnitude of $K \leq K^*(\mathrm{z}) + 2.5$, basing $K^*(\mathrm{z})$ on Lin et al. (2006) and choosing this limit because $2 < K - K^* < 2.5$ is the faintest 0.5mag width bin for which the average K band magnitude error is < 0.1 for all clusters. We convert from apparent K band magnitude to rest-frame luminosity, using a k-correction consistent with Mannucci et al. (2001), and the absolute K band magnitude of the sun, $M_{K,\odot} = 3.39$ (Johnson, 1966). To account for the background we perform this same calculation within 40 apertures on a control field (The UKIDSS-DXS Lockman Hole and XMM-LSS fields, Lawrence et al., 2007), subtracting the average from our total L_K and adding the standard deviations to our L_K errors.

We note that the consistency found in Chapter 3 (Mulroy et al., 2014) between colour-magnitude selected luminosity and spectroscopically confirmed luminosity indicates the accuracy of colour-magnitude member selection in $(J-K)/K$ space.

Richness

We calculate the richness, λ (defined in Rozo et al. 2009 and improved in Rykoff et al. 2012), for the 33 cluster overlap between our sample and the Sloan Digital Sky Survey (SDSS, Gunn et al., 1998; Doi et al., 2010; Alam et al., 2015). This matched filter richness estimator is defined as the sum of the membership probabilities of all the galaxies, which are calculated using the clustercentric radius, magnitude and colour of the galaxies. Its corresponding radius scales with

λ and is not equivalent to an overdensity radius such as r_{500}.

Velocity Dispersion

Finally, we include the velocity dispersion, σ, from Haines et al. (2015), calculated using the biweight scale estimator (Beers et al., 1990) which approaches the standard deviation when the population is Gaussian. The velocity dispersion and its virial radius are determined iteratively, with the radius defined as in Finn et al. (2008) by assuming the cluster is a single isothermal sphere within which the galaxies orbit isotropically.

5.3 Linear Regression

To characterise the scaling relation between observable and mass we linearise the problem by taking the natural log of the values, and perform a linear regression. To do so correctly we have to take into account measurement errors, the mass distribution and the selection criteria. Most commonly used regression methods (e.g. BCES, Akritas & Bershady 1996, and FITEXY, Press et al. 1992; Tremaine et al. 2002) can handle measurement errors, while methods from Kelly (2007) and Mantz (2016) also take into account the mass distribution by modelling it as a Gaussian mixture model inferred from the data.

However the selection function can still introduce significant biases, either directly when the selection variable is on one of the axes, or indirectly due to covariance between this selection variable and the observable of interest (as shown in Section 2.2.1 and quantified for this dataset in Section 5.4.3). It is possible to use these two codes to correct for selection effects when the selection variable is on the dependent axis, by using upper limits and generating 'censored' data below the selection limit in an iterative process. However it is not so straight forward to correct for the bias caused by covariance with the selection variable, i.e. when considering a dependent variable which is not the selection variable.

We therefore develop a multivariate maximum likelihood method similar to Kelly (2007) and Mantz (2016), which simultaneously considers the selection variable alongside all other observables in order to model the covariance and correct for these selection effects.

Table 5.2: Cluster observables

Name	L_X 10^{44}erg/s	k_BT_X keV	$M_{\rm gas}$ $10^{14}M_\odot$	Y_X $10^{14}M_\odot$keV	$Y_{Pl}D_A^2$ 10^{-5}Mpc^2	L_K $10^{12}L_\odot$	λ	σ km/s
Abell2697	$11.68^{+0.44}_{-0.44}$	$6.99^{+0.48}_{-0.38}$	$0.90^{+0.07}_{-0.07}$	$5.42^{+0.40}_{-0.40}$	$10.04^{+0.55}_{-0.55}$	–	$54.75^{+3.77}_{-3.77}$	–
Abell0068	$9.91^{+0.59}_{-0.59}$	$7.66^{+0.77}_{-0.62}$	$0.80^{+0.04}_{-0.04}$	$5.59^{+2.24}_{-2.24}$	$19.20^{+0.66}_{-0.66}$	$12.46^{+2.43}_{-2.71}$	$74.81^{+3.86}_{-3.86}$	1186^{+89}_{-88}
Abell0115	$10.68^{+0.40}_{-0.40}$	$5.93^{+0.39}_{-0.32}$	$0.87^{+0.15}_{-0.15}$	$5.88^{+0.57}_{-0.57}$	$18.38^{+0.55}_{-0.55}$	$14.77^{+2.25}_{-2.58}$	–	1219^{+72}_{-71}
Abell0141	$5.10^{+0.88}_{-0.88}$	$4.78^{+1.34}_{-0.83}$	$0.60^{+0.05}_{-0.05}$	$2.95^{+0.95}_{-0.95}$	$10.66^{+0.52}_{-0.52}$	$15.02^{+1.91}_{-2.08}$	–	–
Abell0209	$13.59^{+1.02}_{-1.02}$	$6.39^{+1.05}_{-0.77}$	$1.44^{+0.08}_{-0.08}$	$8.80^{+1.68}_{-1.68}$	$26.34^{+0.56}_{-0.56}$	$20.51^{+2.56}_{-2.53}$	–	1369^{+65}_{-67}
Abell0267	$6.34^{+2.88}_{-2.88}$	$8.03^{+2.83}_{-1.81}$	$0.70^{+0.05}_{-0.05}$	$7.21^{+2.91}_{-2.91}$	$10.21^{+0.73}_{-0.73}$	$12.71^{+2.50}_{-3.31}$	$64.17^{+3.73}_{-3.73}$	1045^{+52}_{-52}
Abell0291	$3.37^{+0.08}_{-0.08}$	$4.03^{+0.32}_{-0.29}$	$0.47^{+0.04}_{-0.04}$	$1.44^{+0.09}_{-0.09}$	$4.91^{+0.55}_{-0.55}$	$7.79^{+1.34}_{-1.33}$	$46.37^{+2.59}_{-2.59}$	704^{+80}_{-82}
Abell0521	$15.33^{+1.09}_{-1.09}$	$6.72^{+0.33}_{-0.29}$	$1.08^{+0.09}_{-0.09}$	$7.27^{+0.39}_{-0.39}$	$19.07^{+0.68}_{-0.68}$	$14.17^{+2.62}_{-3.09}$	–	–
Abell0586	$6.20^{+0.54}_{-0.54}$	$5.56^{+1.10}_{-0.79}$	$0.73^{+0.06}_{-0.06}$	$3.63^{+0.69}_{-0.69}$	$8.09^{+0.43}_{-0.43}$	$18.30^{+2.43}_{-2.76}$	$68.50^{+3.91}_{-3.91}$	933^{+55}_{-54}
Abell0611	$12.00^{+0.94}_{-0.94}$	$11.96^{+2.50}_{-2.40}$	$0.69^{+0.05}_{-0.05}$	–	$13.53^{+0.83}_{-0.83}$	$13.61^{+3.27}_{-3.40}$	$66.18^{+3.95}_{-3.95}$	1039^{+67}_{-67}
Abell0697	$22.55^{+2.29}_{-2.29}$	$11.06^{+2.16}_{-1.83}$	$1.22^{+0.10}_{-0.10}$	$16.21^{+3.55}_{-3.55}$	$37.00^{+0.75}_{-0.75}$	$13.15^{+3.19}_{-3.19}$	$93.15^{+4.44}_{-4.44}$	1268^{+57}_{-58}
ZwCl0857	$4.50^{+0.19}_{-0.19}$	$3.97^{+0.15}_{-0.46}$	$0.34^{+0.07}_{-0.07}$	$1.40^{+0.11}_{-0.11}$	$1.02^{+0.58}_{-0.58}$	$3.01^{+1.16}_{-1.37}$	$17.21^{+2.27}_{-2.27}$	815^{+79}_{-80}
Abell0750	$2.89^{+0.20}_{-0.20}$	$3.95^{+0.49}_{-0.39}$	$0.55^{+0.06}_{-0.06}$	$2.08^{+0.30}_{-0.30}$	$12.09^{+0.41}_{-0.41}$	$19.73^{+2.73}_{-2.98}$	$98.43^{+4.02}_{-4.02}$	–
Abell0773	$11.11^{+1.14}_{-1.14}$	$7.50^{+1.58}_{-1.12}$	$1.10^{+0.05}_{-0.05}$	$7.46^{+1.39}_{-1.39}$	$20.33^{+0.57}_{-0.57}$	$22.02^{+2.64}_{-2.45}$	$111.21^{+4.52}_{-4.52}$	–
Abell0781	$4.16^{+1.92}_{-1.92}$	$5.92^{+2.40}_{-1.36}$	$0.74^{+0.12}_{-0.12}$	$4.45^{+1.69}_{-1.69}$	$12.05^{+0.84}_{-0.84}$	$16.34^{+3.24}_{-3.24}$	$85.13^{+4.47}_{-4.47}$	–
ZwCl0949	$4.52^{+0.99}_{-0.99}$	$7.31^{+0.94}_{-0.89}$	$0.40^{+0.04}_{-0.04}$	–	$5.05^{+0.45}_{-0.45}$	$7.91^{+1.90}_{-1.93}$	$17.96^{+2.42}_{-2.42}$	–
Abell0901	$1.40^{+0.07}_{-0.07}$	$2.80^{+0.04}_{-0.04}$	$0.26^{+0.03}_{-0.03}$	$0.61^{+0.05}_{-0.05}$	$3.00^{+0.32}_{-0.32}$	$6.94^{+1.58}_{-1.79}$	–	–
Abell0907	$5.91^{+0.22}_{-0.22}$	$5.66^{+0.51}_{-0.41}$	$0.93^{+0.06}_{-0.06}$	$4.01^{+0.33}_{-0.33}$	$15.51^{+0.40}_{-0.40}$	$13.83^{+2.00}_{-2.12}$	–	–
Abell0963	$7.89^{+0.29}_{-0.29}$	$6.53^{+0.62}_{-0.50}$	$0.80^{+0.05}_{-0.05}$	$4.13^{+0.29}_{-0.29}$	$13.24^{+0.49}_{-0.49}$	$13.25^{+2.27}_{-2.46}$	$43.46^{+3.20}_{-3.20}$	1119^{+49}_{-49}
ZwCl1021	$19.66^{+1.47}_{-1.47}$	$9.04^{+1.51}_{-1.13}$	$0.95^{+0.05}_{-0.05}$	$10.80^{+2.50}_{-2.50}$	$13.59^{+0.75}_{-0.75}$	$9.27^{+2.18}_{-2.13}$	$49.33^{+3.51}_{-3.51}$	–
Abell1423	$7.35^{+0.68}_{-0.68}$	$8.20^{+1.54}_{-1.16}$	$0.62^{+0.06}_{-0.06}$	$6.42^{+1.46}_{-1.46}$	$8.59^{+0.43}_{-0.43}$	$9.90^{+2.11}_{-2.26}$	$32.19^{+3.04}_{-3.04}$	–
Abell1451	$6.13^{+1.31}_{-1.31}$	$8.87^{+1.45}_{-1.10}$	$1.02^{+0.05}_{-0.05}$	$7.57^{+1.07}_{-1.07}$	$18.12^{+0.50}_{-0.50}$	$19.17^{+2.49}_{-2.25}$	–	–
ZwCl1231	$7.87^{+0.66}_{-0.66}$	$6.56^{+1.20}_{-0.89}$	$0.69^{+0.11}_{-0.11}$	$5.67^{+1.25}_{-1.25}$	$12.20^{+0.53}_{-0.53}$	$9.43^{+3.14}_{-2.27}$	$43.78^{+3.45}_{-3.45}$	–
Abell1682	$4.99^{+2.00}_{-2.00}$	$6.46^{+2.98}_{-1.49}$	$0.84^{+0.04}_{-0.04}$	$5.18^{+2.36}_{-2.36}$	$12.18^{+0.44}_{-0.44}$	$19.56^{+2.80}_{-3.02}$	$80.00^{+4.10}_{-4.10}$	–
Abell1689	$15.81^{+0.55}_{-0.55}$	$9.71^{+0.64}_{-0.51}$	$1.31^{+0.05}_{-0.05}$	$12.81^{+0.95}_{-0.95}$	$27.77^{+0.48}_{-0.48}$	$23.07^{+2.75}_{-2.92}$	$133.41^{+3.69}_{-3.69}$	1541^{+46}_{-46}
Abell1763	$15.20^{+1.56}_{-1.56}$	$7.67^{+1.64}_{-1.32}$	$1.61^{+0.09}_{-0.09}$	$11.08^{+2.56}_{-2.56}$	$27.90^{+0.53}_{-0.53}$	$21.86^{+4.17}_{-3.72}$	$105.03^{+4.64}_{-4.64}$	1358^{+52}_{-53}
Abell1835	$22.22^{+0.79}_{-0.79}$	$10.16^{+0.68}_{-0.55}$	$1.43^{+0.07}_{-0.07}$	$13.84^{+1.03}_{-1.03}$	$27.95^{+0.79}_{-0.79}$	$21.42^{+3.67}_{-3.33}$	$91.08^{+4.22}_{-4.22}$	1485^{+35}_{-35}
Abell1914	$17.08^{+1.36}_{-1.36}$	$10.06^{+1.47}_{-1.22}$	$1.11^{+0.07}_{-0.07}$	$12.54^{+2.17}_{-2.17}$	$17.69^{+0.34}_{-0.34}$	$13.37^{+1.94}_{-2.16}$	$88.15^{+3.47}_{-3.47}$	1055^{+44}_{-44}
ZwCl1454	$6.66^{+0.27}_{-0.27}$	$4.74^{+0.42}_{-0.34}$	$0.54^{+0.08}_{-0.08}$	$3.10^{+0.36}_{-0.36}$	$8.98^{+0.70}_{-0.70}$	$6.64^{+2.20}_{-2.56}$	$31.02^{+2.76}_{-2.76}$	988^{+84}_{-84}
Abell2009	$6.05^{+0.63}_{-0.63}$	$7.44^{+1.56}_{-1.16}$	$0.69^{+0.05}_{-0.05}$	$4.72^{+1.03}_{-1.03}$	$6.16^{+0.39}_{-0.39}$	$9.08^{+2.39}_{-1.94}$	$65.87^{+3.00}_{-3.00}$	–
RXCJ1504	$16.65^{+1.86}_{-1.86}$	$9.55^{+2.23}_{-1.52}$	$1.06^{+0.08}_{-0.08}$	$11.26^{+3.49}_{-3.49}$	$16.46^{+0.71}_{-0.71}$	$10.31^{+1.84}_{-1.97}$	$42.32^{+3.36}_{-3.36}$	–
Abell2111	$5.93^{+2.76}_{-2.76}$	$7.21^{+2.28}_{-1.52}$	$0.68^{+0.08}_{-0.08}$	$5.49^{+2.12}_{-2.12}$	$11.61^{+0.49}_{-0.49}$	$15.31^{+2.03}_{-2.24}$	$99.06^{+4.36}_{-4.36}$	–
Abell2204	$15.84^{+0.66}_{-0.66}$	$13.38^{+1.15}_{-0.76}$	$1.23^{+0.08}_{-0.08}$	$11.79^{+1.02}_{-1.02}$	$23.69^{+0.44}_{-0.44}$	$19.69^{+2.26}_{-2.12}$	–	–
Abell2219	$32.91^{+2.60}_{-2.60}$	$10.13^{+0.83}_{-0.70}$	$1.68^{+0.11}_{-0.11}$	$17.90^{+1.67}_{-1.67}$	$41.64^{+0.50}_{-0.50}$	$21.72^{+2.62}_{-2.43}$	$107.44^{+4.35}_{-4.35}$	1332^{+59}_{-58}
RXJ1720	$9.63^{+0.57}_{-0.57}$	$7.14^{+0.91}_{-0.73}$	$0.71^{+0.07}_{-0.07}$	$6.60^{+1.00}_{-1.00}$	$12.88^{+0.34}_{-0.34}$	$9.77^{+2.49}_{-1.90}$	$44.89^{+2.85}_{-2.85}$	938^{+37}_{-36}
Abell2261	$13.04^{+1.12}_{-1.12}$	$7.50^{+1.30}_{-1.09}$	$1.23^{+0.06}_{-0.06}$	$8.12^{+1.19}_{-1.19}$	$19.98^{+0.52}_{-0.52}$	$26.60^{+3.32}_{-4.32}$	$84.95^{+4.01}_{-4.01}$	–
RXCJ2102	$4.62^{+0.12}_{-0.12}$	$5.32^{+0.46}_{-0.37}$	$0.46^{+0.05}_{-0.05}$	$2.29^{+0.18}_{-0.18}$	$5.98^{+0.42}_{-0.42}$	$7.77^{+1.46}_{-1.46}$	–	–
RXJ2129	$10.65^{+0.65}_{-0.65}$	$5.94^{+0.75}_{-0.61}$	$0.67^{+0.10}_{-0.10}$	$5.47^{+0.95}_{-0.95}$	$7.21^{+0.57}_{-0.57}$	$7.53^{+2.16}_{-2.31}$	$42.43^{+3.15}_{-3.15}$	879^{+82}_{-82}
Abell2390	$25.43^{+1.16}_{-1.16}$	$10.79^{+0.95}_{-0.84}$	$1.66^{+0.09}_{-0.09}$	$16.91^{+1.57}_{-1.57}$	$40.22^{+0.54}_{-0.54}$	$17.44^{+2.40}_{-2.37}$	$88.68^{+4.50}_{-4.50}$	1372^{+63}_{-62}
Abell2537	$6.63^{+0.72}_{-0.72}$	$9.93^{+3.73}_{-2.44}$	$0.83^{+0.08}_{-0.08}$	$6.30^{+2.30}_{-2.30}$	$13.69^{+0.75}_{-0.75}$	$19.48^{+3.26}_{-3.38}$	$116.80^{+4.65}_{-4.65}$	–
Abell2552	$13.46^{+1.77}_{-1.77}$	$9.69^{+2.75}_{-1.94}$	$1.00^{+0.10}_{-0.10}$	$9.22^{+2.89}_{-2.89}$	$13.78^{+0.74}_{-0.74}$	$19.51^{+4.40}_{-4.78}$	$62.73^{+4.30}_{-4.30}$	–
Abell2631	$14.41^{+1.02}_{-1.02}$	$6.91^{+1.18}_{-0.87}$	$0.97^{+0.12}_{-0.12}$	$6.69^{+1.21}_{-1.21}$	$14.24^{+0.67}_{-0.67}$	$13.75^{+2.68}_{-2.92}$	$79.26^{+4.22}_{-4.22}$	–

5.3.1 Maximum Likelihood Method

We define $\mu \equiv \ln(\text{mass})$ and $\mathbf{S}$ as the vector of all observables for a given cluster $S \equiv \ln(\text{observable})$.

The probability that there exists a cluster with given observables and mass is

$$P(\mathbf{S}, \mu|\boldsymbol{\theta}, \psi) = P(\mathbf{S}|\mu, \boldsymbol{\theta})P(\mu|\psi), \tag{5.1}$$

where $\boldsymbol{\theta}$ characterises the scaling relation between observable and mass, and ψ characterises the mass distribution. Simulations have shown that weak-lensing measurements are on average an unbiased measure of true mass when the fitting radius is carefully considered, with a scatter of $\sigma_{\ln M_{\rm WL}|M_{\rm True}} = 0.2$ (e.g. Oguri & Hamana, 2011; Becker & Kravtsov, 2011; Bahé et al., 2012). As such we use a mass function as parameterised in Evrard et al. (2014), convolved with a Gaussian to account for this scatter.

$P(\mathbf{S}|\mu, \boldsymbol{\theta})$ is the probability of a vector of observables given a mass and scaling relations between observables and mass:

$$P(\mathbf{S}|\mu, \boldsymbol{\theta}) \propto \exp\left\{-\frac{1}{2}(\mathbf{S} - \boldsymbol{b}\mu - \mathbf{a})^T\Sigma^{-1}(\mathbf{S} - \boldsymbol{b}\mu - \mathbf{a})\right\}, \tag{5.2}$$

where $\boldsymbol{\theta} = (\mathbf{a}, \boldsymbol{b}, \Sigma)$, the intercept, slope and covariance matrix of observables. The one-dimensional (i.e. single observable) equivalent is

$$P(S|\mu, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(S - b\mu - a)^2}{2\sigma^2}\right\}, \tag{5.3}$$

where $\theta = (a, b, \sigma)$, with σ the intrinsic scatter in the relation. We continue with single observable notation for simplicity.

In reality we cannot observe these true values of S; our measurements, S_o, have observational uncertainties. We include these by marginalising over all possible values of true observable S:

$$P(S_o, \mu|, \theta, \psi) = \int \mathrm{d}S\ P(S_o|S)\ P(S, \mu|\theta, \psi) \tag{5.4}$$

$$= \int \mathrm{d}S\ P(S_o|S)\ P(S|\mu, \theta)\ P(\mu|\psi). \tag{5.5}$$

Assuming these measurement errors are log-normal gives

$$P(S_o|S) = \frac{1}{\sqrt{2\pi\sigma_{\rm err}^2}} \exp\left\{-\frac{(S_o - S)^2}{2\sigma_{\rm err}^2}\right\}. \tag{5.6}$$

Table 5.3: Scaling relation parameters predicted by self-similarity

Observable S	E(z) factor for S, f_S	E(z) factor for μ, f_μ	Self-Similar Slope
L_X	-1	1	1.33
$k_B T_X$	0	1	0.66
$M_{\rm gas}$	0	0	1
Y_X	1	1	1.66
$Y_{\rm Pl} D_A^2$	1	1	1.66
L_K	0	0	1
λ	0	0	1
σ	0	1	0.33

We make the same log-normal assumption and include errors on the mass measurements μ_o in the same way to obtain $P(S_o, \mu_o|, \theta, \psi)$.

Finally, and most significantly, we are able to include the effect of the selection function, because this method considers a vector of observables, including the selection observable. Our selection function is simply a redshift dependent $L_{X,\rm RASS}$ threshold, which is taken into account using a redshift dependent step function.

5.3.2 Self-Similar Scaling

On the scale of galaxy clusters the dominant force is gravity, which is scale invariant. This means that, according to self-similarity (Kaiser, 1986), galaxy clusters are simply scaled versions of each other, with their properties determined only by their mass and redshift. The mass determines the strength of the gravitational potential, while the redshift determines the evolution of the critical density.

From the simple assumptions of self-similarity we can predict how cluster observables scale with mass, which we derived in Section 2.2.2 and summarise in Table 5.3. We divide the masses by a pivot ($M_P = 7.34 \times 10^{14} M_\odot$) and fit the following scaling relations:

$$S + f_S \epsilon = b_S(\mu + f_\mu \epsilon) + \ln a_S \tag{5.7}$$

where $\epsilon = \ln E(z)$ and f_μ and f_S are the factors of ϵ associated with mass and each observable respectively, shown in Table 5.3.

Table 5.4: Scaling relation parameters

Observable S	Intercept a	Slope b	Scatter $\sigma_{S\|\mu}$
L_X	$8.44^{+0.58}_{-0.59}$	$1.32^{+0.20}_{-0.20}$	$0.39^{+0.05}_{-0.05}$
$k_B T_X$	$7.10^{+0.32}_{-0.32}$	$0.63^{+0.13}_{-0.13}$	$0.22^{+0.04}_{-0.04}$
M_{gas}	$0.97^{+0.04}_{-0.04}$	$0.89^{+0.11}_{-0.11}$	$0.15^{+0.03}_{-0.04}$
Y_X	$6.27^{+0.47}_{-0.46}$	$1.62^{+0.21}_{-0.21}$	$0.36^{+0.06}_{-0.06}$
$Y_{\rm Pl} D_A^2$	$15.87^{+0.90}_{-0.90}$	$1.35^{+0.17}_{-0.17}$	$0.26^{+0.04}_{-0.04}$
L_K	$16.52^{+0.64}_{-0.64}$	$0.72^{+0.11}_{-0.11}$	$0.08^{+0.07}_{-0.06}$
λ	$78.72^{+4.95}_{-4.91}$	$0.78^{+0.18}_{-0.18}$	$0.29^{+0.05}_{-0.05}$
σ	$1117.81^{+23.1}_{-23.19}$	$0.43^{+0.05}_{-0.05}$	$0.02^{+0.01}_{-0.01}$

5.4 Results

5.4.1 Scaling Relations

The scaling relation parameters resulting from applying the method described in Section 5.3 to the data described in Section 5.2 are summarised in Table 5.4 and shown in Figure 5.2.

The results of the X-ray scaling relations are in good agreement with the self-similar predictions (although $M_{\rm gas}$ has a $\sim 1.1\sigma$ disagreement). This tells us that the behaviour of the ICM in galaxy clusters is well described by self-similarity, with clusters forming through spherical collapse.

As Y_X and $Y_{\rm Pl}$ measurements both trace the thermal energy of the ICM, we might expect them to have the same scaling relation with mass. We find however that while Y_X scales consistently with self-similarity, $Y_{\rm Pl}$ disagrees by almost 2σ. A limitation of the Planck observations are their angular resolution of $\sim 7'$ ($\gtrsim r_{500}$ for clusters at this redshift), which require the flux to be measured by fitting a profile out to $5r_{500}$ and interpolating within r_{500}. This process has been shown to have minimal bias in simulations (Planck Collaboration et al., 2016c), however substructures were not included and could lead to biases in the measurements. In future work we will investigate further whether the Planck resolution is the cause of this discrepancy by analysing $Y_{\rm SZA}$ measured from Sunyaev Zel'dovich Array (SZA) data, which has a high enough angular resolution that we don't need to assume a profile.

A simple model predicts that L_K scales linearly with M, but our results are flatter by 2.3σ. This requires that either or both of the cluster stellar mass or the cluster stellar mass-to-light

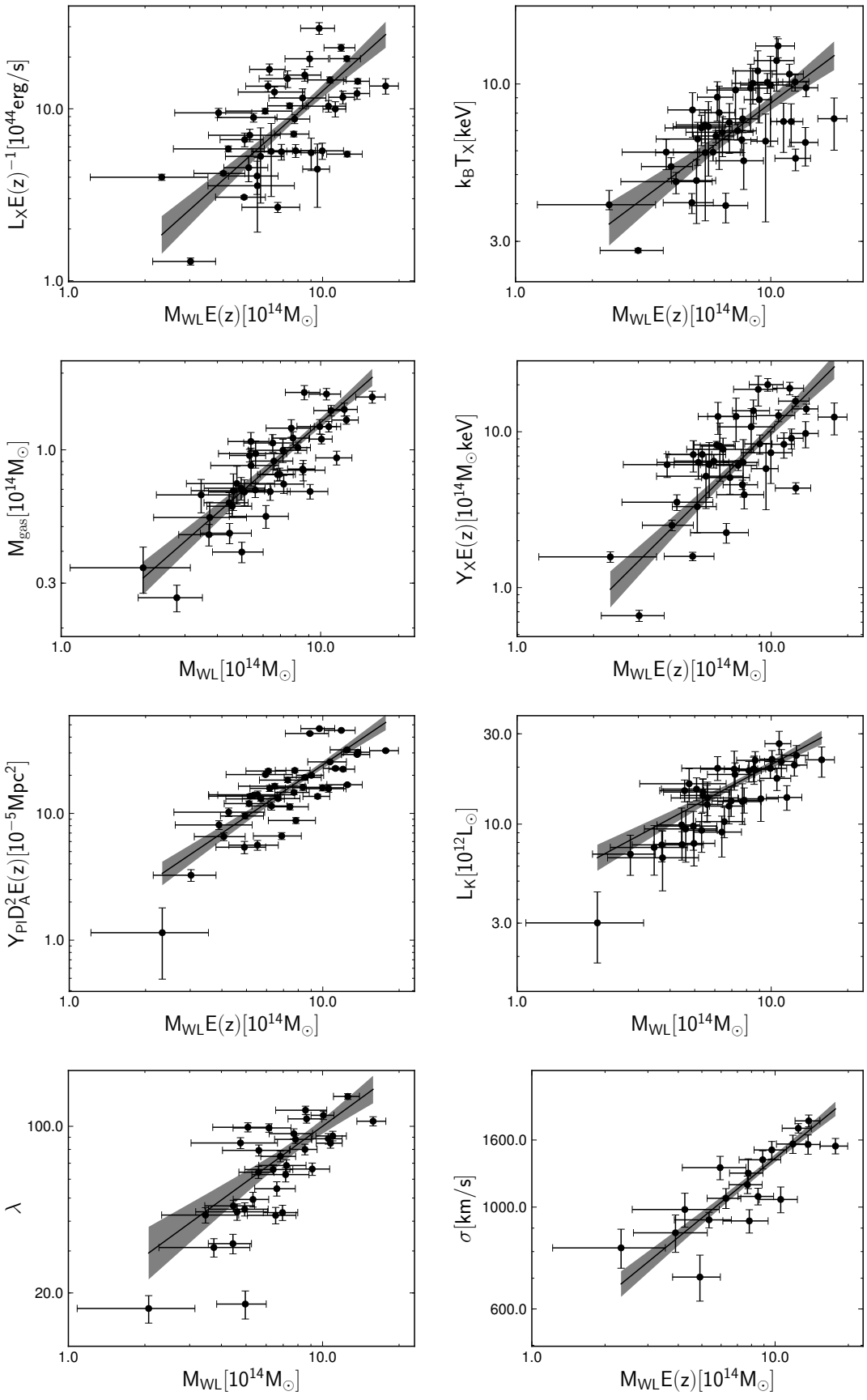

Figure 5.2: Scaling relations between cluster observables and $M_{\rm WL}$, where we show the data points, resultant scaling relation and 68 per cent confidence region.

ratio are dependent on cluster mass. Similarly the simple model predicts λ to scale linearly. This prediction assumes that each galaxy has the same mass, an assumption known not to be correct and seen as such by the disagreement in slope on the order of 1.2σ. The slope of the σ relation is steeper than predicted by self-similar, by $\sim 2.5\sigma$, which suggests that clusters are not simply spherical virialised systems, perhaps due to substructure.

We highlight the low scatter of some of the observables, in particular σ, L_K and $M_{\rm gas}$. This makes them promising mass proxies for future surveys, in particular L_K because its measurement requires only near-infrared imaging and minimal analysis, and is therefore an appealing observable for large surveys with wide field data.

There is a large literature on scaling relations (for a recent review see Giodini et al., 2013), however comparisons are complicated by the variety of measurement and fitting methods. Here we focus on the only other paper in the literature to take a multivariate approach to modelling the selection effect. Mantz et al. (2016b) consider similar X-ray parameters, and in agreement with our results they find consistency with self-similarity. They find their non-core-excised L_X slope slightly higher than self-similar, but conclude that this is due to departures from self-similarity in the inner ($< 0.15 r_{500}$) region of the cluster. $M_{\rm gas}$ is the only X-ray observable in our analysis that includes the core, perhaps explaining the slight deviation from self-similarity in this parameter.

We find agreement with the values of intrinsic scatter found by Mantz et al. (2016b) for L_X and T_X. Our Y_X scatter is higher than their result by $\sim 2.3\sigma$, but we note that this work sets an upper limit of 0.11 on the scatter for $M_{\rm gas}$, while we allow this parameter to be free and find 0.16 ± 0.04, which goes some way to solve the discrepancy.

We note that the scatter between weak-lensing mass and true mass is a free parameter in the model of Mantz et al. (2016b), who constrain it to 0.17 ± 0.06, consistent with the assumption of 0.2 that we make in Section 5.3.1.

5.4.2 Covariance

The maximum likelihood analysis constrains the covariance between the intrinsic scatter with mass of two observables. A given element of a covariance matrix can range between $\pm\sigma_{S_i|\mu}\sigma_{S_j|\mu}$, while the correlation coefficient, $r_{S_iS_j}$, is equal to the covariance divided by $\sigma_{S_i|\mu}\sigma_{S_j|\mu}$, and is

Table 5.5: Correlation coefficients

	L_X	k_BT_X	$M_{\rm gas}$	Y_X	$Y_{\rm Pl}D_A^2$	L_K	λ
k_BT_X	$0.58^{+0.13}_{-0.13}$	–	–	–	–	–	–
M_{gas}	$0.79^{+0.10}_{-0.10}$	$0.25^{+0.22}_{-0.22}$	–	–	–	–	–
Y_X	$0.91^{+0.05}_{-0.05}$	$0.77^{+0.10}_{-0.10}$	$0.65^{+0.15}_{-0.14}$	–	–	–	–
$Y_{\rm Pl}D_A^2$	$0.47^{+0.16}_{-0.16}$	$0.17^{+0.21}_{-0.21}$	$0.49^{+0.17}_{-0.17}$	$0.52^{+0.15}_{-0.15}$	–	–	–
L_K	$-0.41^{+0.36}_{-0.36}$	$-0.33^{+0.37}_{-0.40}$	$-0.48^{+0.34}_{-0.37}$	$-0.37^{+0.36}_{-0.37}$	$-0.06^{+0.40}_{-0.42}$	–	–
λ	$-0.26^{+0.17}_{-0.17}$	$-0.23^{+0.23}_{-0.22}$	$-0.11^{+0.21}_{-0.21}$	$-0.16^{+0.19}_{-0.18}$	$0.01^{+0.20}_{-0.20}$	$0.48^{+0.36}_{-0.36}$	–
σ	$0.04^{+0.56}_{-0.53}$	$-0.15^{+0.58}_{-0.56}$	$-0.18^{+0.56}_{-0.56}$	$-0.03^{+0.55}_{-0.54}$	$0.06^{+0.57}_{-0.61}$	$-0.01^{+0.65}_{-0.65}$	$-0.09^{+0.59}_{-0.59}$

therefore limited to the range ± 1. A positive r tells us that on average the objects with positive residuals (i.e. that lie above the scaling relation) in observable S_1 also have positive residuals in observable S_2, and the same for negative residuals. While a negative r tells us that on average objects with positive residuals in S_1 have negative residuals in S_2, and vice versa. The magnitude of r tells us how strongly these residuals are related. These correlation coefficients are summarised in Table 5.5 and shown in Figure 5.3.

Among the strongest features of these results is positive covariance between any pair of observables that probe the gas content of the cluster (X-ray parameters and $Y_{\rm Pl}$). There is also positive covariance between the observables that probe the stellar content of the cluster (L_K and λ). Interestingly there is negative covariance between any pair of gas and stellar observable, consistent with the 'closed box' view of clusters. In this picture, the content of a cluster is representative of the content of the Universe as a whole, and as such has a fixed baryon fraction. Thus the higher the gas fraction, the lower the stellar fraction, and vice versa. This anti-correlation is consistent with the hydrodynamical simulations of Wu et al. (2015), who found $r_{M_{\rm gas},M_\star} = -0.72 \pm 0.02$.

The correlation coefficients for velocity dispersion are all consistent with zero but are very weakly constrained, perhaps due to its intrinsic scatter with mass being consistent with zero.

There are a few previous observational constraints on covariance between X-ray parameters. Our results are better constrained but in agreement with $r_{L_X,M_{\rm gas}}$ and $r_{T_X,M_{\rm gas}}$ measured by Maughan (2014), and broadly consistent with their r_{L_X,T_X} value. While inconsistent with earlier work (Mantz et al., 2010, 2015), our r_{L_X,T_X} value is consistent both in value and in level of constraint with the updated work of Mantz et al. (2016b). We find mostly disagreement between

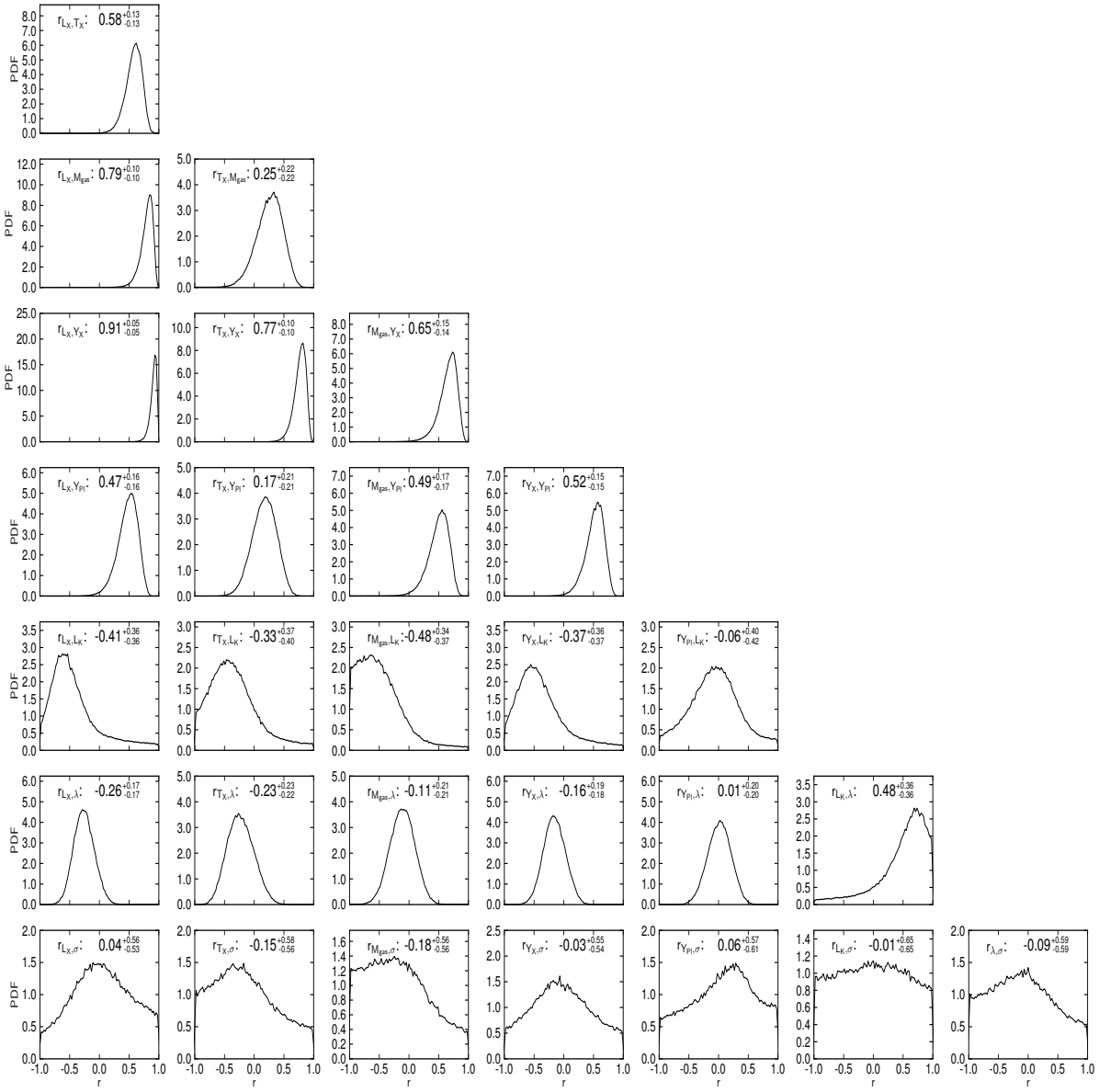

Figure 5.3: Probability distribution function (PDF) of the correlation coefficients between pairs of observables at fixed mass.

Table 5.6: Scaling relation parameters without correction for selection effects

Observable S	Intercept a	Slope b	Scatter $\sigma_{S\|\mu}$
L_X	$7.74^{+0.69}_{-0.74}$	$1.18^{+0.27}_{-0.28}$	$0.50^{+0.07}_{-0.08}$
$k_B T_X$	$6.88^{+0.32}_{-0.34}$	$0.64^{+0.13}_{-0.14}$	$0.23^{+0.04}_{-0.05}$
M_{gas}	$0.99^{+0.04}_{-0.04}$	$1.02^{+0.13}_{-0.14}$	$0.17^{+0.05}_{-0.05}$
Y_X	$6.10^{+0.59}_{-0.65}$	$1.52^{+0.28}_{-0.29}$	$0.48^{+0.08}_{-0.09}$
$Y_{\rm Pl} D_A^2$	$13.98^{+0.99}_{-1.02}$	$1.47^{+0.20}_{-0.22}$	$0.32^{+0.06}_{-0.07}$
L_K	$16.63^{+0.74}_{-0.73}$	$0.92^{+0.13}_{-0.14}$	$0.11^{+0.05}_{-0.06}$
λ	$75.66^{+4.60}_{-4.76}$	$1.26^{+0.20}_{-0.22}$	$0.24^{+0.07}_{-0.07}$
σ	$1182.35^{+48.68}_{-46.59}$	$0.48^{+0.09}_{-0.10}$	$0.09^{+0.03}_{-0.04}$

The scaling relation parameters fitted using the Kelly (2007) method, with no correction for selection effects. Comparison with the scaling relation fitted using our maximum likelihood method which corrects for selection effects, shown in Table 5.4, quantifies the bias of these effects and the importance of correcting for them.

our results and those of Mantz et al. (2016a), who studied a sample of 40 relaxed clusters. However, as demonstrated by Mantz et al. (2016b), this dynamically selected sample represents only a limited fraction of the residual phase space. Interestingly, we find agreement with their $r_{L_X,M_{\rm gas}}$ value, perhaps demonstrating that T_X is the variable most affected by limiting analysis to a relaxed sample.

5.4.3 Selection Bias

Table 5.6 shows the scaling relation parameters fitted using the Kelly (2007) method, with no correction for selection effects. Comparison with the scaling relation fitted using our maximum likelihood method which corrects for selection effects, shown in Table 5.4, quantifies the bias of these effects and the importance of accounting for them.

5.5 Summary

In this chapter we have introduced a new multivariate Bayesian approach to linear regression, which accounts for the underlying distribution of clusters and the effects of a selection function. Using this method we have performed the first simultaneous analysis of galaxy cluster scaling relations between observables across the electromagnetic spectrum and weak-lensing mass. We are able to provide good constraints on all scaling relations, and on the correlation coefficients between each pair of observables at fixed mass.

We find consistency with self-similar predictions for X-ray parameters, suggesting that the ICM behaviour is as predicted for self-similarity, i.e. that clusters form via a single spherical collapse and the dominant force is gravity. We find disagreement between Y_X and $Y_{\rm Pl}$, but note that this may be caused by the low angular resolution of Planck observations, and plan to investigate this with comparison with SZA data in the future. The probes of the stellar content (L_K and λ) however are both shallower than self-similar. We note the small intrinsic scatter of σ, L_K and $M_{\rm gas}$, and highlight that in particular L_K is easy to measure from survey data and so is a promising mass proxy for future surveys.

The covariance between any two measures of the gas content is positive, as is the covariance between L_K and σ, which measure the stellar content. We also find negative correlation between any pair of gas observable and stellar observable, consistent with the closed box picture of galaxy clusters, in which the gas and stellar fraction of a cluster can vary but its total baryon fraction is constant.

Overall these results suggest that clusters are formed via the spherical collapse model consistent with self-similarity, and maintain a constant baryon fraction during their evolution.

Chapter 6

Conclusions

In this thesis I have presented work on the integrated observables of galaxy clusters and their relations with each other and with measures of the level of disturbance in the cluster.

In Chapter 1 we motivated the importance of galaxy cluster mass measurements for constraining cosmological parameters. In Chapter 2 we considered galaxy cluster observables, how they can be related to cluster masses, and what they tell us about astrophysics.

In Chapter 3 we confirmed the near-infrared luminosity of a cluster as a promising mass proxy for future wide field surveys, both because of its low intrinsic scatter with mass and because it can be easily measured from shallow survey data.

In Chapter 4 we showed that the results of Chapter 3 are not restricted to that particular wavelength as might have been expected, as we found low intrinsic scatter when considering luminosities in bluer bands. We also presented trends between cluster colour and the level of disturbance in a cluster, suggesting that the total cluster star formation rates are more varied in undisturbed clusters than in disturbed clusters.

In Chapter 5 we introduced a new linear regression method to correct for selection effects, and through our multivariate analysis were able to place competitive constraints on the scaling relation parameters and correlation coefficients of cluster observables across the electromagnetic spectrum, from microwave through to X-ray. Our results are consistent with self-similarity and the closed box picture of galaxy clusters.

Future Work

A next step for this work is to incorporate measurements of $Y_{\rm SZA}$ from the SZA into our linear regression analysis. This will help us to interpret the disagreement we find between the Y_X and $Y_{\rm Pl}$ scaling relations. If $Y_{\rm SZA}$ is consistent with Y_X this suggests the discrepancy is due to the limited angular resolution of the Planck data, while consistency with $Y_{\rm Pl}$ suggests a more

fundamental difference between SZ and X-ray measurements.

Another next step is to consider deviations from the scaling relations of Chapter 5. Deviations by individual clusters can tell us about the physics of that particular cluster, while trends between these deviations and a measure of the state of the cluster can highlight the underlying physics causing intrinsic scatter in the scaling relations.

The negative covariance between any observable probing the gas content and any observable probing the stellar content motivates future work on the stellar and gas fractions of clusters. Doing this properly, by taking into account not only the galaxies but also the intracluster light, will show how close the baryon fraction in clusters is to the Universal baryon fraction, and whether there is a mass dependence. The quality of the LoCuSS data will allow us to look at this as a function of radius to reveal any variation in the baryonic composition, and perhaps if there exists a typical radius within which clusters are truly representative of the Universe as a whole.

References

Abbott B. P., et al., 2016, Physical Review Letters, 116, 061102

Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706

Alam S., et al., 2015, APJS, 219, 12

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARAA, 49, 409

Andersson K., et al., 2011, ApJ, 738, 48

Andreon S., Hurn M. A., 2010, MNRAS, 404, 1922

Applegate D. E., et al., 2014, MNRAS, 439, 48

Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&AP, 474, L37

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&AP, 517, A92

Bahé Y. M., McCarthy I. G., King L. J., 2012, MNRAS, 421, 1073

Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681

Balogh M. L., McGee S. L., 2010, MNRAS, 402, L59

Balogh M. L., Mazzotta P., Bower R. G., Eke V., Bourdin H., Lu T., Theuns T., 2011, MNRAS, 412, 947

Bardeau S., Soucail G., Kneib J.-P., Czoske O., Ebeling H., Hudelot P., Smail I., Smith G. P., 2007, A&AP, 470, 449

Bartalucci I., Arnaud M., Pratt G. W., Démoclès J., van der Burg R. F. J., Mazzotta P., 2017, A&AP, 598, A61

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Bell E. F., de Jong R. S., 2001, ApJ, 550, 212

Blanton M. R., Roweis S., 2007, AJ, 133, 734

Boehringer H., Hensler G., 1989, A&AP, 215, 147

Böhringer H., et al., 2004, A&AP, 425, 367

Bonamente M., Joy M., LaRoque S. J., Carlstrom J. E., Nagai D., Marrone D. P., 2008, ApJ, 675, 106

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Budzynski J. M., Koposov S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, MNRAS, 423, 104

Burns J. O., 1998, Science, 280, 400

Caldwell R. R., Dave R., Steinhardt P. J., 1998, Physical Review Letters, 80, 1582

Carlberg R. G., Yee H. K. C., Ellingson E., 1997a, ApJ, 478, 462

Carlberg R. G., et al., 1997b, ApJL, 485, L13

Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARAA, 40, 643

Castignani G., Benoist C., 2016, A&AP, 595, A111

Cavagnolo K. W., Donahue M., Voit G. M., Sun M., 2008, ApJL, 683, L107

Cavagnolo K. W., Donahue M., Voit G. M., Sun M., 2009, APJS, 182, 12

Chilingarian I. V., Melchior A.-L., Zolotukhin I. Y., 2010, MNRAS, 405, 1409

Chilingarian I. V., Zolotukhin I. Y., 2012, MNRAS, 419, 1727

Clowe D., Luppino G. A., Kaiser N., Gioia I. M., 2000, ApJ, 539, 540

Cole S., et al., 2001, MNRAS, 326, 255

Dariush A. A., Raychaudhury S., Ponman T. J., Khosroshahi H. G., Benson A. J., Bower R. G., Pearce F., 2010, MNRAS, 405, 1873

Deason A. J., Conroy C., Wetzel A. R., Tinker J. L., 2013, ApJ, 777, 154

Doi M., et al., 2010, AJ, 139, 1628

Dressler A., Shectman S. A., 1988, AJ, 95, 985

Dyson F. W., Eddington A. S., Davidson C., 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881

Ebeling H., Edge A. C., Allen S. W., Crawford C. S., Fabian A. C., Huchra J. P., 2000, MNRAS, 318, 333

Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562

Finn R. A., Balogh M. L., Zaritsky D., Miller C. J., Nichol R. C., 2008, ApJ, 679, 279

Fitzpatrick E. L., 1999, PASP, 111, 63

Frieman J. A., Turner M. S., Huterer D., 2008, ARAA, 46, 385

Garn T., Best P. N., 2010, MNRAS, 409, 421

Giodini S., et al., 2009, ApJ, 703, 982

Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, SSR, 177, 247

Girardi M., Borgani S., Giuricin G., Mardirossian F., Mezzetti M., 2000, ApJ, 530, 62

Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147

Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, ApJ, 778, 14

Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1

Gunn J. E., et al., 1998, AJ, 116, 3040

Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111

Haines C. P., et al., 2009, ApJ, 704, 126

Haines C. P., et al., 2013, ApJ, 775, 126

Haines C. P., et al., 2015, ApJ, 806, 101

Hansen S. M., McKay T. A., Wechsler R. H., Annis J., Sheldon E. S., Kimball A., 2005, ApJ, 633, 122

Hitomi Collaboration et al., 2016, Nature, 535, 117

Hjorth J., Oukbir J., van Kampen E., 1998, MNRAS, 298, L1

Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, MNRAS, 427, 1298

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Ilbert O., et al., 2009, ApJ, 690, 1236

Jee M. J., et al., 2011, ApJ, 737, 59

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Johnson H. L., 1966, ARAA, 4, 193

Kaiser N., 1986, MNRAS, 222, 323

Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460

Kauffmann G., Charlot S., 1998, MNRAS, 297, L23

Kelly B. C., 2007, ApJ, 665, 1489

Kneib J.-P., et al., 2003, ApJ, 598, 804

Kochanek C. S., White M., Huchra J., Macri L., Jarrett T. H., Schneider S. E., Mader J., 2003, ApJ, 585, 161

Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128

Kravtsov A. V., Borgani S., 2012, ARAA, 50, 353

Kravtsov A., Vikhlinin A., Meshscheryakov A., 2014, preprint, (arXiv:1401.7329)

Laganá T. F., Zhang Y.-Y., Reiprich T. H., Schneider P., 2011, ApJ, 743, 13

Lawrence A., et al., 2007, MNRAS, 379, 1599

Leauthaud A., et al., 2012, ApJ, 746, 95

Lin Y.-T., Mohr J. J., Stanford S. A., 2003, ApJ, 591, 749

Lin Y.-T., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745

Lin Y.-T., Mohr J. J., Gonzalez A. H., Stanford S. A., 2006, ApJL, 650, L99

Lu T., Gilbank D. G., Balogh M. L., Bognat A., 2009, MNRAS, 399, 1858

Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116

Mancone C. L., Gonzalez A. H., 2012, PASP, 124, 606

Mannucci F., Basile F., Poggianti B. M., Cimatti A., Daddi E., Pozzetti L., Vanzi L., 2001, MNRAS, 326, 745

Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, MNRAS, 406, 1773

Mantz A. B., et al., 2015, MNRAS, 446, 2205

Mantz A. B., 2016, MNRAS, 457, 1279

Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016a, MNRAS, 456, 4020

Mantz A. B., et al., 2016b, MNRAS, 463, 3582

Marrone D. P., et al., 2012, ApJ, 754, 119

Martino R., Mazzotta P., Bourdin H., Smith G. P., Bartalucci I., Marrone D. P., Finoguenov A., Okabe N., 2014, MNRAS, 443, 2342

Maughan B. J., 2014, MNRAS, 437, 1171

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

McDonald M., Veilleux S., Rupke D. S. N., Mushotzky R., 2010, ApJ, 721, 1262

McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937

McGee S. L., Balogh M. L., Wilman D. J., Bower R. G., Mulchaey J. S., Parker L. C., Oemler A., 2011, MNRAS, 413, 996

Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&AP, 514, A93

Messier C., 1781, Technical report, Catalogue des Nébuleuses des amas d'Étoiles (Catalog of Nebulae and Star Clusters)

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Mulroy S. L., et al., 2014, MNRAS, 443, 3309

Muzzin A., Yee H. K. C., Hall P. B., Lin H., 2007, ApJ, 663, 150

Muzzin A., et al., 2012, ApJ, 746, 188

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Nelan J. E., Smith R. J., Hudson M. J., Wegner G. A., Lucey J. R., Moore S. A. W., Quinney S. J., Suntzeff N. B., 2005, ApJ, 632, 137

Oguri M., Hamana T., 2011, MNRAS, 414, 1851

Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2010a, PASJ, 62, 811

Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010b, ApJ, 721, 875

Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, ApJL, 769, L35

Okabe N., Smith G. P., 2016, MNRAS, 461, 3794

Pasquali A., Gallazzi A., Fontanot F., van den Bosch F. C., De Lucia G., Mo H. J., Yang X., 2010, MNRAS, 407, 937

Pearson R. J., Ponman T. J., Norberg P., Robotham A. S. G., Farr W. M., 2015, MNRAS, 449, 3082

Perlmutter S., et al., 1999, ApJ, 517, 565

Planck Collaboration et al., 2013, A&AP, 550, A129

Planck Collaboration et al., 2016a, A&AP, 594, A8

Planck Collaboration et al., 2016b, A&AP, 594, A13

Planck Collaboration et al., 2016c, A&AP, 594, A27

Poole G. B., Babul A., McCarthy I. G., Sanderson A. J. R., Fardal M. A., 2008, MNRAS, 391, 1163

Popesso P., Biviano A., Böhringer H., Romaniello M., Voges W., 2005, A&AP, 433, 431

Pranger F., Böhm A., Ferrari C., Maurogordato S., Benoist C., Höller H., Schindler S., 2014, A&AP, 570, A40

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&AP, 498, 361

Press W. H., Schechter P., 1974, ApJ, 187, 425

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing

Ramella M., Boschin W., Geller M. J., Mahdavi A., Rines K., 2004, AJ, 128, 2022

Rasia E., et al., 2012, New Journal of Physics, 14, 055018

Ratra B., Peebles P. J. E., 1988, Phys. Rev. D, 37, 3406

Rawle T. D., et al., 2014, MNRAS, 442, 196

Richard J., et al., 2010, MNRAS, 404, 325

Riess A. G., et al., 1998, AJ, 116, 1009

Rines K., Geller M. J., Kurtz M. J., Diaferio A., Jarrett T. H., Huchra J. P., 2001, ApJL, 561, L41

Rines K., Geller M. J., Diaferio A., Kurtz M. J., Jarrett T. H., 2004, AJ, 128, 1078

Rossetti M., Molendi S., 2010, A&AP, 510, A83

Rozo E., et al., 2009, ApJ, 699, 768

Ruel J., et al., 2014, ApJ, 792, 45

Rykoff E. S., et al., 2012, ApJ, 746, 178

Sanderson A. J. R., Edge A. C., Smith G. P., 2009, MNRAS, 398, 1698

Santos J. S., Rosati P., Tozzi P., Böhringer H., Ettori S., Bignamini A., 2008, A&AP, 483, 35

Sarazin C. L., 1988, X-ray emission from clusters of galaxies

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Smail I., Ellis R. S., Dressler A., Couch W. J., Oemler Jr. A., Sharples R. M., Butcher H., 1997, ApJ, 479, 70

Smith G. P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417

Smith R. J., Lucey J. R., Price J., Hudson M. J., Phillipps S., 2012, MNRAS, 419, 3167

Springel V., et al., 2005, Nature, 435, 629

Stroe A., et al., 2015, MNRAS, 450, 646

Sutherland R. S., Dopita M. A., 1993, APJS, 88, 253

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878

Tremaine S., et al., 2002, ApJ, 574, 740

Umetsu K., Broadhurst T., 2008, ApJ, 684, 177

Umetsu K., et al., 2009, ApJ, 694, 1643

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Vikhlinin A., et al., 2009, ApJ, 692, 1033

Weinberg S., 1989, Reviews of Modern Physics, 61, 1

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232

Wu H.-Y., Evrard A. E., Hahn O., Martizzi D., Teyssier R., Wechsler R. H., 2015, MNRAS, 452, 1982

Ziparo F., et al., 2016, A&AP, 592, A9

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

van Haarlem M. P., Wise M. W., Gunst A. W., others 2013, A&AP, 556, A2

van Weeren R. J., Röttgering H. J. A., Brüggen M., Hoeft M., 2010, Science, 330, 347

von der Linden A., Wild V., Kauffmann G., White S. D. M., Weinmann S., 2010, MNRAS, 404, 1231